

This Page Is Inserted by IFW Operations
and is not a part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

**As rescanning documents *will not* correct images,
please do not report the images to the
Image Problems Mailbox.**

THIS PAGE BLANK (USPTO)

8B05

SEARCH REQUEST FORM

8-70

Requestor's

Name:

S. Ugar

Serial

Number:

08/949,904

Date:

8/3/98

Phone:

305-2181

Art Unit:

1642

Search Topic:

Please write a detailed statement of search topic. Describe specifically as possible the subject matter to be searched. Define any terms that may have a special meaning. Give examples or relevant citations, authors, keywords, etc., if known. For sequences, please attach a copy of the sequence. You may include a copy of the broadest and/or most relevant claim(s).

Please Search SEQ ID #1

SEQ ID #2

SEQ ID #3

Reverse Transcribed SEQ ID #2

also Interference Search

Push For Seq ID #1
" " #2
" " #3

*08-848439
08-796153

Reverse Transcribed SEQ ID #2

I need this by COB 8/5

Thanks

Susan

STAFF USE ONLY

Date completed:

8-5-98

Searcher:

BCE 2-4291

Terminal time:

3

Elapsed time:

4-15

CPU time:

Total time:

Number of Searches:

Number of Databases:

11

Search Site

STIC

CM-1

Pre-S

Type of Search

2 N.A. Sequence

6 A.A. Sequence

Structure

Bibliographic

Vendors

✓ IG

STN

Dialog

APS

Geninfo

SDC

DARC/Questel

MPI Other

THIS PAGE BLANK (USPTO)


```
*****
[MPERCH]
[RESEARCH]
[UNIT]
*****
Release 3.1A John F. Collins, Blocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPERCH_nn n.a. - n.a. database search, using Smith-Waterman algorithm

Run on: Wed Aug 5 04:44:53 1998; MasPar time 2726.89 Seconds
1365.563 Million cell updates/sec

Tabular output not generated.

Title: >US-08-848-439-1
Description: (1-2027) from US08848439.seq
Perfect Score: 2027
N.A. Sequence: 1 GAATTCGGGCTTCATGGCCT.....AAAAAAGGCGCGC 2027
Comp: CTTAAGCCGGAGTACCGGA.....TTTTTTTTTCGGCGGG

Scoring table: TABLE default
Gap 6

Nmatch STD : Dbase 0; Query 0

Searched: 532261 seqs, 918536377 bases x 2

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: emb155
1:em_ba 2:em_htg 3:em_hum1 4:em_hum2 5:em_in 6:em_om
7:em_or 8:em_ov 9:em_pat 10:em_pl 11:em_ro
genbank107
12:gb_ba 13:gb_htg 14:gb_in 15:gb_om 16:gb_ov 17:gb_pat
18:gb_ph 19:gb_pl 20:gb_pr1 21:gb_pr2 22:gb_ro 23:gb_st
24:gb_sts 25:gb_sy 26:gb_um 27:gb_v1

Statistics: Mean 12.042; Variance 6.972; scale 1.727

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No. Score Match Length DB ID Description Pred. No.
-----
1 1043 51.5 2028 22 MM088567 Mus musculus secreted 0.00e+00
2 993 49.0 2031 22 AF017989 Mus musculus secreted 0.00e+00
3 959 47.3 1776 22 D50462 Mouse SDF5 mRNA, compl 0.00e+00
4 765 37.8 882 21 AF017986 Homo sapiens secreted 0.00e+00
5 110 5.4 1984 21 AF017988 Homo sapiens secreted 3.62e-50
6 106 5.2 2094 21 AF017987 Homo sapiens secreted 1.75e-47
7 106 5.2 4240 15 BT085945 Bos taurus frizzled-re 1.75e-47
8 103 5.1 2075 21 AF001900 Homo sapiens secreted 1.77e-45
9 103 5.1 4469 21 AF056087 Homo sapiens secreted 1.77e-45
10 101 5.0 2659 22 MM088566 Mus musculus secreted 3.80e-44
11 97 4.8 7218 17 I66494 Sequence 14 from paten 1.72e-41
12 96 4.7 7218 17 I66494 Sequence 14 from paten 7.88e-41
13 83 4.1 2075 16 AF006508 Gallus gallus crescent 2.54e-32
14 60 3.0 2184 21 HSU82169 Human frizzled homolog 7.84e-18
15 58 2.9 2334 20 HSU43318 Human putative transme 1.28e-16
C
```

```
16 2.8 2328 16 AF031830 Gallus gallus 7-transm 5.11e-16
17 2.8 2421 22 MM043321 Mus musculus putative 2.03e-15
18 2.5 1881 22 AF054623 Mus musculus frizzled- 1.84e-12
19 4.8 2.4 1181 16 AF059570 Xenopus laevis secreta 1.00e-10
20 4.9 2.4 10772 14 AF012089 Drosophila melanogaste 2.67e-11
21 4.6 2.3 10772 14 AF012089 Drosophila melanogaste 1.38e-09
22 4.5 2.2 1912 22 RATERZ2H Rattus norvegicus Dros 5.08e-09
23 4.2 2.1 1727 16 AF031831 Gallus gallus 7-transm 2.36e-07
24 4.1 2.0 1923 21 HUMER12 Human frizzled gene pr 8.32e-07
25 3.9 2.5 215 17 I28278 Sequence 5 from patent 9.95e-06
26 3.8 1.9 4540 22 RATERZ2H Rattus norvegicus Dros 3.39e-05
27 3.1 1.5 215 17 I28278 Sequence 5 from patent 1.18e-01
28 3.0 1.5 354 14 OF089259 Oxytricha fallax 57kd 3.53e-01
29 3.0 1.5 4297 19 ATHSC701 A.thaliana hsc70-1 gen 3.53e-01
30 3.1 1.5 38262 14 CELK02A2 Caenorhabditis elegans 1.18e-01
31 3.1 1.5 149435 13 HS111189 Human DNA sequence *** 1.18e-01
32 2.9 1.4 156 24 D366D7S D. melanogaster SFS de 1.03e+00
33 2.9 1.4 307 24 AF010529 Plasmodium falciparum 1.03e+00
34 2.8 1.4 354 14 OF089259 Oxytricha fallax 57kd 2.96e+00
35 2.8 1.4 397 24 G14542 human SFS SHG-11394. 2.96e+00
36 2.9 1.4 1073 19 ATCOR47 A.thaliana cor47 mRNA. 1.03e+00
37 2.8 1.4 1566 17 E03096 DNA encoding calphobin 2.96e+00
38 2.9 1.4 2028 21 HSU56813 Human polycystin mRNA 1.03e+00
39 2.9 1.4 2112 22 MM088567 Mus musculus mRNA for EB 1.03e+00
40 2.9 1.4 2260 22 MM043320 Mus musculus putative 1.03e+00
41 2.9 1.4 4798 14 DFU14576 Dictyostellum discoide 1.03e+00
42 2.9 1.4 8595 14 DD014576 Arabidopsis thaliana g 2.96e+00
43 2.8 1.4 72698 19 AB007645 Arabidopsis thaliana D 2.96e+00
44 2.8 1.4 110102 19 ATF28J12 Arabidopsis thaliana D 2.96e+00
45 2.9 1.4 118710 13 AC004392 *** SEQUENCING IN PROG 1.03e+00

ALIGNMENTS

RESULT 1 MM088567 2028 bp mRNA ROD 22-APR-1997
LOCUS Mus musculus secreted frizzled related protein sFRP-2 (Sfrp2) mRNA,
DEFINITION complete cds.
ACCESSION U88567
NID 91946342
KEYWORDS house mouse.
SOURCE Mus musculus
ORGANISM Eukaryotes; Mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae;
Murinae; Mus.
REFERENCE 1 (bases 1 to 2028)
AUTHORS Rattner,A., Hsieh,J.-C., Smallwood,P.M., Gilbert,D.J.,
Copeland,N.G., Jenkins,N.A. and Nathans,J.
TITLE A family of secreted proteins contains homology to the
cysteine-rich ligand-binding domain of frizzled receptors
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 94 (7), 2859-2863 (1997)
MEDLINE 97250455
REFERENCE 2 (bases 1 to 2028)
AUTHORS Rattner,A., Hsieh,J.-C., Smallwood,P.M., Gilbert,D.J.,
Copeland,N.G., Jenkins,N.A. and Nathans,J.
TITLE Direct Submission
JOURNAL Submitted (05-FEB-1997) Molecular Biology and Genetics, HHMI/Johns
Hopkins University School of Medicine, 725 North Wolfe Street, PCTB
805, Baltimore, MD 21205
FEATURES
source
1. 2028
/organism="Mus musculus"
/db_xref="taxon:10090"
/chromosome="3"
/map="near Fgg"
236..1123
/gene="Sfrp2"
236..1123
/gene="Sfrp2"
/gene="Sfrp2"
/codon_start=1
/product="secreted frizzled related protein sFRP-2"
/db_xref="PID:g1946343"
```

[illegible]

[illegible]

QY	567	CGATGACCTAGACGAGACCATCCAGCCATGCCACCTCGCTCGCTGCGAGGTGAAGGACCG	626
Db	625	CTGCGCCCGGGTCATCTCGCCCTTCGGGTTCCTCCCTGGCGACACATGCTGGAGTGCAGCCG	684
QY	627	CTGCGCCCGGGTCATCTCGCCCTTCGGGTTCCTCCCTGGCGACACATGCTGGAGTGCAGCCG	685
Db	685	TTTCCCGCAGGACACGAGCCTCTGATCCCTCCCTCGCTAGTAGCGACACACCTCTCTGCCGGC	744
QY	687	TTTCCCGCAGGACACGAGCCTTTGGATCCCTCCCTCGCTAGTAGCGACACACCTCTCTGCCAGC	746
Db	745	CACAGAGGAGCTCCCAAGGTGTGGAAGCTTGCACCAACCAAGATGAGGACGACCAACGA	804
QY	747	CACCGAGGAGCTCCCAAGGTATGTGAAGCTTGCACCAATAAAAATGATGATGACCAACGA	806
Db	805	CATCATGGAACCCCTTTGTAAAAATGACTTCGCACCTGAAAATCAAAGTGAAGGAGATAAC	864
QY	807	CATATGGAACGCCTTTGTAAAAATGATTTGCACTGAAAATAAAAGTGAAGGAGATAAC	866
Db	865	GTACATCAACAGACACCAAGATCATCTCTGGAGACAAAGAGCAAGACCATTTTCAAGCT	924
QY	867	CTACATCAACCGAGATACCAAAATCATCTGGAGACCAAGAGCAAGACCATTTTCAAGCT	926
Db	925	GAAGCGGTGTCGGAAGGAGCCTGAAGAAATCCGTGCTGTGGCTCAAGACAGCCTCGCA	984
QY	927	GAAGGTGTGTCGGAAGGAGCCTGAAGAAATCGGTGCTGTGGCTCAAGACAGCTTGCA	986
Db	985	GTGCACCTGTGAGGAGATCAAGCAGATCAACGCTCCGTATCTTGGTCATGGGACAGAACA	1044
QY	987	GTGCACCTGTGAGGAGATCAAGCAGATCAACGCGCCATCTGGTCATGGGACAGAACA	1046
Db	1045	GGCGGCGAACTGGTGATCACTCCGTGAAACGGTGGCAGAGGCGCAGAGAGATTCAA	1104
QY	1047	GGGTGGGAGCTGGTGATCACCTCGGTGAAGCGGTGGCAGAGGCGCAGAGAGATTCAA	1106
Db	1105	GCACATCTCCGACAGATCCGAAGCTGCAATGCTAGTTTCCCACTGGGTGGCTTCTCT	1166
QY	1107	GCACATCTCCGACAGATCCGAAGCTGCAATGCTAGTTTCCCACTGGGTGGCTTCTCT	1168
Db	1165	CCATCCAGGCGCTGAGCTCTGTAGACCACTTGC - CTCGG - ACCTCATTTCGGTTCCTCC	1222
QY	1167	CAGCGCTGCTCAGACAGCAGCTGACCATTTCTGCTCGGGATCTCAGCTCCGCTTCC	1226
Db	1223	AAGCAGACTCCGGAAAGCTACAGCCCCAGCTTGGAGCGCTTGCCCTGCCTCTGCATG	1288
QY	1227	AAGCAGACTCTAGCT - GCTCCAGTCTCAGCTGGGAGCTTCCCGCTGCCTTTTGACG	1290
Db	1283	TGTGTATCCCTAATGCTCCTGAGTTTAAAGCCCTAGAGGCGCTTGGAAACCCATAGCT	1342
QY	1286	TTTGTGATCCCAAGCATTTCTGAGTTTAAAGGCCACAGGAG - - - TGG - - - TAGCT	1395
Db	1343	GTATTTACGGAAGCGAAAGCCATCCAGATCTTGTACAATATTTCAAACCTAATAAAT	1402
QY	1336	GTATTTACCTAAG - GAAAGGCCACCCGAATCTTGTAGAAATATTTCAACATATAAAT	1394
Db	1403	CATGACTATTTTATGAAGTTTTTAGAACAGCTCGTTTAAAGTTAGTTTGAATAGCTGT	1465
QY	1395	CATGAATATTTTATGAAGTTTAAAAATAGCTCACTTTTAAAGCTAGTTTGAATAGGTGC	1454
Db	1463	AGTACTTTGACCGAGGGGCATTTTCTCTCTTTGGTCAGTCTGTTGGCTTATACCGTGCA	1522
QY	1455	A - - ACTGTGACTTTGGTCTGTGGTGTGTTGTTTGT - GT - T - TTGAGTCAAG - C - TGAT	1506
Db	1523	CTTAGTGTGCAATGTCAGGCGAATGTTTCTTTTTTTTTTTTTTTTTTTTTTCCCTCTGTGGTCT	1584
QY	1507	TTTCACTTCCAC - TGAGTGTGCTAATACATGCAATTTGCTTCAATTTTCTCTGTGGGCC	1555
Db	1583	AAGCTTTGGGTCCAGACTTAGTTGAGATAAGCTGGCTGTTATCTCAAAAGCTTCTCTC	1642
QY	1566	AACTTTGGGTCAACAACCCGTTTGAGATAAAGCTGGCTGTTATCTCAACATCTTCTATC	1625
Db	1643	AGTTTCAGGCTGAGATTCGGCATCTAAGTCTTTCAACATTTTGGTGTGCTGTTTANGGCC	1702

Db	909	GTGGGGTGGCTTCTCTCATCCAGCGCCTGAGCTCTGTAGTACCACTTCCGCTCCGCGACC	968
Qy	1151	ATCTGATGGCTCGACAGGCTGCTCCAGAGCAGCGGTGACCATTTCTGCTCCGGGATC	1210
Db	969	TCATTTCGGGTTTCCCAAGCACAGTCCGGGAAGACTACAGCCCGAGCTTGGAGCGCTTG	1028
Qy	1211	TCAGTCCCGTTCCCAAGCACACTCTAGCT-GCTCAGCTCAGCTGGCAGCTTCC	1269
Db	1029	CCCTGCCCTCTGCATGTGTATCCCTAACATGCTCCTGAGTTATAAGGCCCTTAGGAGCC	1088
Qy	1270	CCCTGCCCTTTTGCAGGTTTGCATCCCGAGCATTTCTCAGTTATAAGGCCACAGAG---	1326
Db	1089	TTGAACCCCATAGCTGTTTTCACGGAAGCGAAAGCCCATCCAGATCTTGTACAATA	1148
Qy	1327	-TGGA-----TAGCTGTTTTCACCTAAG-GAAAGCCACCCGAATCTTGTAGAATA	1378
Db	1149	TTCAAACTAATAAATCATGACTATTTTATGAAGTTTATAGACAGCTGTTTAAAGGT	1208
Qy	1379	TTCAAACTAATAAATCATGAATATTTTATGAAGTTTAAATAGCTCACTTAAAGCT	1438
Db	1209	AGTTTGAATAGCTGTAGTACTTTGACCCGAGGGGCATTTTCTCTTTTGGTCAGTCTGT	1268
Qy	1439	AGTTTGAATAGGTGCA--ACTGTGACTTGGCTCTGGTTGTTGTTGTTGTT-T-T	1493
Db	1269	TGGCTTATACCGTGCATTTAGTTGGCCATGTGACGGGAATTTGTTCTTTTTTTTTTT	1328
Qy	1494	TGAGTCAG--C-TGATTTTCTACTTCCAC-TGAGGTTGTCTATAACATCAAAATTCCTCA	1549
Db	1329	TTTCCCTCTGGTCTAGCTTGTGGTCCGAGACTTGTAGTTAGATAAGCTGGCTGTTA	1388
Qy	1550	ATTTTCTCTGTGGGCCAAACTGTGGGTGCACAAACCCCTGTTGAGATAAAGCTGGCTGTTA	1609
Db	1389	TCTCAAGCTCTCTCAGTCCAGCTCAGAACTCGCATCTAAAGTCTTCAACACATTTTCT	1448
Qy	1610	TCTCAACATCTTTCATCAGCTCCAGCTGAGACTCAGTCTCTAAGCTTACAACTTCAAT	1669
Db	1449	TGCTCGTTTTATGCCCTCATGAGCTCTGACCATTCATGATTCATGCTTCCCATCCAGTACAGA	1508
Qy	1670	CATTTTATACCTTCAATGG-GAACTTAACTGTTACATGTAT-CACATTCAGCTACAAAT	1727
Db	1509	ACTTCAGTTTATAAGCA-CACAGTAACCATTCCTCATTCATGATGCCCTCAAAATAAAA	1567
Qy	1728	ACTTCCATTTATAGAAGCAGCATTAACCATTTCT-ATAGCATGATTTCTTCAAGTAAAG	1786
Db	1568	GTGAATACAGTCTATAAATTCACAGAGTATTTTAAGCTTTGTTTAAACATCTTTTAATTC	1627
Qy	1787	GCAA-AGA-TATA-AATTTTATAATGACTTGAG--TACTTTAAGCTTGTTTAARA-C	1840
Db	1628	AATTTTAAATCATTTTTTTTGCAAACTAAATCATATGTAGCTTACCTGTATAATACGTAG	1687
Qy	1841	A-TTCTTACTTAATTTTGCA-AATTAAGCCCATTTAGCTTACCTGTATAATACATAG	1898
Db	1688	TAGTTGACCTGGAAAGTTGTAANAATATTCTTTAAACGACACTGTAANAATTTTCAGAT	1747
Qy	1899	TAGTTTACCTTTAAAGTTGTAANAATATTCTTTAAACCAACACTGTAANAATTTTCAGAT	1958
Db	1748	AAACATTAATTTCTTTGTATATAAATTT 1776	
Qy	1959	AAACATTAATTTCTT-GTATATAAATTT 1986	
RESULT	4	AF017986 882 bp mRNA PRI 21-SEP-1997	
LOCUS		Homo sapiens secreted apoptosis related protein 1 (SARPL) mRNA,	
DEFINITION		partial cds.	
ACCESSION		AF017986	
NID		g2415414	
KEYWORDS			
SOURCE		human.	
ORGANISM		Homo sapiens	
		Eukaryotes; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;	
		Primates; Catarrhini; Homnidae; Homo.	
REFERENCE		1 (bases 1 to 882)	

AUTHORS	Melkonyan, H., Chang, W.C., Shapiro, J.P., Mahadevappa, M., Fitzpatrick, P.A., Kiefer, M.C., Tomei, D.B. and Umansky, S.R.
TITLE	SARPs - a new family of proteins that regulate apoptosis
JOURNAL	Unpublished
REFERENCE	2 (bases 1 to 882)
AUTHORS	Melkonyan, H., Prochazka, V. and Umansky, S.R.
TITLE	Direct Submission
JOURNAL	Submitted (11-AUG-1997) LXR Biotechnology, Inc., 1401 Marina Way South, Richmond, CA 94804, USA
FEATURES	Location/Qualifiers
Source	1..882
	/organism="Homo sapiens"
	/db_xref="taxon:9606"
gene	1..>882
	/gene="SARPL"
CDS	265..>882
	/gene="SARPL"
	/note="similar to frizzled-like receptor CRD"
	/codon_start=1
	/product="secreted apoptosis related protein 1"
	/db_xref="PID:g2415415"
	/translation="MLQGGPSLLLLFLASHCCCLGSARGLFLFGQPDFSYKRSNCKPIP ANLQCHGEYONMRLPNLLGHETKEVLEQACAWIPLYNKQCHPDTKFLCLSLFAPV CLNLDDETTQPCHSRCVQVQKDRCAPYMSAFPDPDMLCEDRFPQDNLDLCIPLASSDHL PATFEAPKVCCEAKNKNDNDIMETLCKNDFALKIKRREITINR"
BASE COUNT	167 a 318 c 233 g 163 t 1 others
ORIGIN	
Query Match	37.8%; Score 766; DB 21; Length 882;
Best Local Similarity	97.6%; Pred. No. 0.00e+00;
Matches	841; Conservative 0; Mismatches 10; Indels 11; Gaps 4;
Db	29 CTGCTCATTTCTGCTCCCGGGTCCGAGCCCGCGAGCTGCGCGGGGCTTCAGCGC 88
Qy	19 CTACTCATTTCTGCTCCCGGGTCCGAGCCCGCGAGCTGCGCGGGGCTTCAGCGC 78
Db	89 CTCGCCCGGCTGCTCCCGGGTCCCGGTTCTTCCGCGCCCGAGCCCGGGTTCGACG 148
Qy	79 CTCGCCCGGCT--CCTCCCGGTTCTCCGCTTCTCCGCGCCCGAGCCCGGCTGCCAGC 136
Db	149 TTTTCGGGGCCCCAGTTCGACCCAGAGAGAGCGCGCCCGGAGACAGCTCGAATCC 208
Qy	137 TTTTCGGGGCCCCAGTTCGACCCAGAGAGAGCGCGCCCGGAGCAAGCTCGAATCC 196
Db	209 GCGCGCTTCGCGCTTAAACAGCTCCG-TCCCTCTACCCCTTAGGGGTGCGC--CCACGA 265
Qy	197 GCGCGCTTCGCGCTTCCCGGGTCCGCTCCCTCTGCCCCCTCGGGGTGCGCGCCAGCA 256
Db	266 TGTGTCAGGGCCCTGGGTGCGTGTGCTGCTTCTTCCGCTCGCCTCGACGTGCTGGGCT 325
Qy	257 TGTGTCAGGGCCCTGGGTGCGTGTGCTGCTTCTTCTCCGCTCGCACTGCTGCTGGGCT 316
Db	326 CGGCGCGGGGCTTCTTCTTTTGGCCAGCCCGAGCTTCTCTTACAGCGAGCAATTCGA 385
Qy	317 CGGCGCGGGGCTTCTTCTTTTGGCCAGCCCGAGCTTCTTCTTACAGCGAGCAATTCGA 376
Db	386 AGCCCATCCCGGCCAACCTGCAGCTGTGCCAGCGCATCGAATACAGAACATCGGGTGC 445
Qy	377 AGCCCATCCCGGCCAACCTGCAGCTGTGCCAGCGCATCGAATACAGAACATCGGGTGC 436
Db	446 CCAACCTGCTGGGCCACGAGACCATGAAGAGGTGCTGGAGCAGCGCGCGCTTGGATTC 505
Qy	437 CCAACCTGCTGGGCCACGAGACCATGAAGAGGTGCTGGAGCAGCGCGCGCTTGGATTC 496
Db	506 CGCTGTCATGAAGCAGTGGCCCGGACACCAAGAGTTCTTGTGCTGCTCTTCGCTTCGCCC 565
Qy	497 CGCTGTCATGAAGCAGTGGCCCGGACACCAAGAGTTCTTGTGCTGCTCTTCGCTTCGCCC 556
Db	566 CGCTGCTGCTGATGACTTAGACGAGACCATTCAGCCATCGCTCGNTGCGTGCAGG 625
Qy	557 CCGTCTGCTGATGACTTAGACGAGACCATTCAGCCATCGCTCGCTGCTGCTGCTGCAGG 616

```

Db 626 TGAAGGATCGCTGCGCCCGGTCATGTCGCCCTC-----CCTGGCCCGACATGCTTG 679
QY 617 TGAAGGACCGCTGCGCCCGGTCATGTCGCCCTTCCGCTTCCGCTGCGCCGACATGCTTG 676
Db 680 AGTGGACCGCTTCCCGCAGGACAAGACCTTTCGATCCCGCTCGCTAGCAGGACACCC 739
QY 677 AGTGGACCGCTTCCCGCAGGACAAGACCTTTCGATCCCGCTCGCTAGCAGGACACCC 736
Db 740 TCCTGCCAGCAGCAGGAGAGCTCCAAAGTATGTGAAGCTCCAAATATAAAATGATG 799
QY 737 TCTGCCAGCAGCAGGAGAGCTCCAAAGTATGTGAAGCTCCAAATATAAAATGATG 796
Db 800 ATGACAACGACATAATGGAACCGCTTTGTAATAATGATTTGCACTGAAATAAAGTGA 859
QY 797 ATGACAACGACATAATGGAACCGCTTTGTAATAATGATTTGCACTGAAATAAAGTGA 856
Db 860 AGGAGATACTACATCAACCG 881
QY 857 AGGAGATACTACATCAACCG 878

RESULT 5
LOCUS AF017988 1984 bp mRNA PRI 21-SEP-1997
DEFINITION Homo sapiens secreted apoptosis related protein 3 (SARP3) mRNA,
complete cds.
ACCESSION AF017988
NID 92415418
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
REFERENCE
1 (bases 1 to 1984)
AUTHORS Melkonian,H., Chang,W.C., Shapiro,J.P., Mahadevappa,M.,
Fitapatric,P.A., Kiefer,M.C., Tomei,D.L. and Umansky,S.R.
TITLE SARPs - a new family of proteins that regulate apoptosis
JOURNAL Unpublished
REFERENCE
2 (bases 1 to 1984)
AUTHORS Melkonian,H., Prochazka,V. and Umansky,S.R.
JOURNAL Direct Submission
TITLE Submitted (11-AUG-1997) LXR Biotechnology, Inc., 1401 Marina Way
JOURNAL South, Richmond, CA 94804, USA
FEATURES
source
Location/Qualifiers
1..1984
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="15"
/tissue_type="pancreas"
1..1984
/gene="SARP3"
216..1169
/gene="SARP3"
/feature="SARP3"
/note="similar to frizzled-like receptor, extracellular
protein"
/codon_start=1
/product="secreted apoptosis related protein 3"
/db_xref="PID:g2415419"
/translation="MRAAAGGVTAALALLGALHWAPCEYDYVQAEPLHG
RYSRPPQCLDIPADLPICHTVGYKMRPLNLEHSAEYVQQAASSWPLAKRCHS
DTVOELSLFAPVLDRIPIPCRSCEAVRAGCAPLMEAYGFPPWEMLHKHFPJLND
LCIAVOFGLPATAFPVTKICAOEHSADGLMOMCSDFVVKMRIRKTIENGDR
KLIGAKKKLKLKLPGLPKRDKTKRLVLMKNGAGGCPQLDLSLFLVMGRKVDGL
LLMAYVRMDKKNKMKFAVKFMSYPCSLYFFFYGAAPHP"
BASE COUNT 353 a 634 c 627 g 370 t
ORIGIN
Query Match 5.4%; Score 110; DB 21; Length 1984;
Best Local Similarity 73.0%; Pred. No. 3.62e-50;
Matches 246; Conservative 0; Mismatches 82; Indels 9; Gaps 3;

Db 380 CATCCCTGCGACCTGCGCTCTGCCACAGGTTGGCTACAAAGCGATGCGGCTGCCCAA 439
QY 381 CATCCCGGCCAACCTGCAGCTGTGCCACGCGATCGAATACCAACATGCGGCTGCCCAA 440

```

```

Db 440 CCTGCTGGACGACGAGACCTGCGCGAAGTGAAGCAGCAGGACGAGCTGGCTGCCGCT 499
QY 441 CCTGCTGGCCACGAGACCATGAAGAGGTGCTGGAGAGGCGCGCTTGGATCCCGCT 500
Db 500 GCTGGCCAAAGCCTGCGACTCGGATACGAGGTCTTCCTGCTGCTCTTTTGGCGCCGT 559
QY 501 GGTCTATGAGACGAGTCCACCGGACACCAAGAAGTTCCTGCTGCTCTTTCGCCCGCT 560
Db 560 CTGCTCTCGA-----CG-GCCCATCTACCCGCTCGCTGCTGGAGCGGTGG 610
QY 561 CTGCTCTGATGACCTAGAGAGACCATCCAGCCACTCGCTCTGCTGCTGAGTGA 620
Db 611 CGCCGCGCTGCGCGCGCTCATGAGGCGCTACGCGCTTCCCTGCGCTGAGATGCTGACTG 670
QY 621 GGACCGCTGCGCGCGCTCATGCTGCGCTTCCGCTTCCCTGCGCGCGACATGCTTGAGTG 680
Db 671 CCACAAGTTCCCGCTGGACACGACCTCTGCAATCGCC 707
QY 681 CGACCGTTTCCCGCGAGGACAGCGCTTTCATCCCG 717

RESULT 6
LOCUS AF017987 2094 bp mRNA PRI 21-SEP-1997
DEFINITION Homo sapiens secreted apoptosis related protein 2 (SARP2) mRNA,
complete cds.
ACCESSION AF017987
NID 92415416
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
REFERENCE
1 (bases 1 to 2094)
AUTHORS Melkonian,H., Prochazka,V., Chang,W.C. and Umansky,S.R.
TITLE SARPs - a new family of proteins that regulate apoptosis
JOURNAL Unpublished
REFERENCE
2 (bases 1 to 2094)
AUTHORS Melkonian,H., Prochazka,V. and Umansky,S.R.
JOURNAL Direct Submission
TITLE Submitted (11-AUG-1997) LXR Biotechnology, Inc., 1401 Marina Way
JOURNAL South, Richmond, CA 94804, USA
FEATURES
source
Location/Qualifiers
1..2094
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="8"
/tissue_type="heart"
1..2094
/gene="SARP2"
303..1247
/gene="SARP2"
/note="similar to frizzled-like receptor, extracellular
protein"
/codon_start=1
/product="secreted apoptosis related protein 2"
/db_xref="PID:g2415417"
/translation="MGIRSGGRRGALGVLLALGAALLAVGSAEYDYVFSFSDIG
PYQSGRYTKPQCVDIPADLRICHNVYKKMVLNLEHETMAEVKQOASWVPLLN
KNCRAGQVFLCSLEAFVCLDRPIYPCRWLCEAVRDSCEPWFQFGYWPMLKDKF
PEGDVKTAMPPTPEASKPGQITVPCPDNKLSEALIEHLCASEFALRWKIKEVK
EGDKYIVPKKPKLKLGLPIKKDLKLVLYLKNAGADPCPCHOLDNLSHHFLIMGRKVK
SOYLTAIHKWDKKMKFKFMKMKKHECPTFSQVFK"
BASE COUNT 484 a 598 c 631 g 381 t
ORIGIN
Query Match 5.2%; Score 106; DB 21; Length 2094;
Best Local Similarity 71.8%; Pred. No. 1.75e-47;
Matches 249; Conservative 0; Mismatches 89; Indels 9; Gaps 5;

Db 452 CTCTACACCAAGCAGCCTCAGTGGTGGACATCCCGCGGACCTCGGCTGTGCCCAA 511
QY 351 CTCTCTACACGCGCAGCANTTGAAGCCCATCCCGCGCACCTGTCGACCG 410

```

```
Db 512 CGTGGCTACAAGAAGATGCTGCTGCCCAACCTGCTGGAGCAGAGACCATTGCGGAGGT 571
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 411 CATGAATACCAAGAACATCGGCTGCCCAACCTGCTGGCCACGAGACCATTGAAGAGGT 470
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 572 GAACGACAGCCAGCAGCTGGTGGCCCTGCTCAACAAGAACTGCCACGCCGCCACCA 631
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 471 GCTGGAGCAGCCGCGCTTGGATCCCGCTGGTCAATGAAGCAGTGCACCCGGACACAA 530
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 632 GGTCTTCTGCTGCTCTTCGCGCCGCTGCTGCT-G--GACCG-G-C---CCATCTA 682
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 531 GAGTTCTGCTGCTCTTCGCGCCGCTGCTGCTGATGACCTAGACGAGACCACCA 590
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 683 CCGTGTGCTGCTGCTGCGAGCGCTGCGGACTCTGCGAGCGGCTCATGCACTTCT 742
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 591 GCCATGCCACTGCTGCTGCGAGTGAAGGACCGTGCGCCCGCTCATGTCGCGCT 650
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 743 CGGCTTCTACTGCGCCGAGATGCTTAAGTGTGACAAGTTTCCCGGAGG 789
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 651 CGGCTTCCCTGGCCGACATGCTGATGCGACCGTTTCCCGCAGG 697

RESULT 7
LOCUS BTU85945 4240 bp mRNA MAM 21-AUG-1997
DEFINITION Bos taurus frizzled-related protein Frza mRNA, complete cds.
ACCESSION U85945
NID 92337936
KEYWORDS Bos taurus.
SOURCE Bos taurus.
ORGANISM Bos taurus.
REFERENCE 1 (bases 1 to 4240)
AUTHORS Duplaa,C. and D'Amore,P.A.
TITLE Identification and cloning of a novel secreted form of mammalian frizzled: evidence to suggest a role in the control of growth and differentiation
JOURNAL Unpublished
AUTHORS Duplaa,C. and D'Amore,P.A.
TITLE Direct Submission
JOURNAL Submitted (20-JAN-1997) INSERM U441, AV du Haut Leveque, Pessac 33600, France
FEATURES
    source
        1..4240
            /organism="Bos taurus"
            /db_xref="taxon:9913"
            /cell_type="endothelial"
            /tissue_type="heart"
            198..1124
            /note="frizzled related protein"
            /codon_start=1
            /product="Frza"
            /db_xref="PID:g2337937"
            /translation="MGGRWAAAGALIALAAGLAAGSASEYDVVSFQSDIGAYQSGR
            FTKPQCVDPADLRCHNVGKRVLPNLEHETMAEVKQOASWVPLLNKCHIG
            TOVFLSLFAPVCLDRPIYPCRWLCEAVRDSCEPVQFEGFYPEMLKCDKRPQGDVC
            IATPNATEASPGQTVCPDCELKSEAIIEHLCASEFALRMKIKVEKRENGDKK
            IVPKKRKLPLKPKIKKELKLVLYLKGADCPCHOLDNLSHHFLIMGRKRVKSYLLT
            AIHKWDKKNKEFTFMKMKNHCEPTFSQVFK"
            210..263
            sig_peptide
            misc_feature
            966..968
            /note="encodes N-glycosylation site"
BASE COUNT 1015 a 1073 c 1112 g 1040 t
ORIGIN
Query Match 5.2%; Score 106; DB 15; Length 4240;
Best Local Similarity 71.4%; Pred. No. 1.75e-47;
Matches 252; Conservative 0; Mismatches 92; Indels 9; Gaps 5;

Db 329 CTTCTACCAAGCACCAGCAGTCGCTGGACATCCCGGACACCTCGGCTGCCACAA 388
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
```

```
QY 351 CTTCTCTACAAGCGGAGCAATTGCAAGCCCATCCCGGCCAACCTGCAGCTGTGCCACGG 410
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 389 CGTGGCTACAAGAGGATGCTGCTGCCCAACCTGCTGGAGCAGAGACCATTGCGCGAGGT 448
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 411 CATGAATACCAAGAACATCGGCTGCCCAACCTGCTGGCCACGAGACCATTGAAGAGGT 470
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 449 GAACGACAGCCAGCAGCTGGTGGCCCTGCTCAACAAGAACTGCCACATCGGCACCCA 508
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 471 GCTGGAGCAGCCGCGCTTGGATCCCGCTGGTCAATGAAGCAGTGCACCCGGACACAA 530
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 509 GGTCTTCTGCTGCTCTTCGCGCCGCTGCTGCT-G--GACCG-G-C---CCATCTA 559
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 531 GAGTTCTGCTGCTCTTCGCGCCGCTGCTGCTGATGACCTAGACGAGACCACCA 590
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 560 CCGTGTGCTGCTGCTGCGAGCGCTGCGGACTCTGCGAGCGGCTCATGCACTTCT 619
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 591 GCCATGCCACTGCTGCTGCGAGTGAAGGACCGTGCGCCCGCTCATGTCGCGCT 650
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 620 CGGCTTCTACTGCGCCGAGATGCTCAAGTCGACAAAGTTTCCCGAGGCGACG 672
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 651 CGGCTTCCCTGGCCGACATGCTGATGCGACCGTTTCCCGCAGGACAACG 703

RESULT 8
LOCUS AF001900 2075 bp mRNA PRI 25-JUN-1997
DEFINITION Homo sapiens secreted frizzled-related protein mRNA, complete cds.
ACCESSION AF001900
NID 92213818
KEYWORDS human.
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
REFERENCE 1 (bases 1 to 2075)
AUTHORS Finch,P.W., He,X., Kelley,M.J., Uren,A., Schaudties,R.P.,
Popescu,N.C., Rudikoff,S., Aaronson,S.A., Varmus,H.E. and
Rubin,J.S.
TITLE Purification and Molecular Cloning of a Secreted, Frizzled-Related Antagonist of Wnt Action
JOURNAL Proc. Natl. Acad. Sci. U.S.A. (1997) In press
REFERENCE 2 (bases 1 to 2075)
AUTHORS Finch,P.W., Aaronson,S.A. and Rubin,J.S.
TITLE Direct Submission
JOURNAL Submitted (30-APR-1997) LCMB, DBS/NCI, 9000 Rockville Pike, Bethesda, MD 20892, USA
REFERENCE 3 (bases 1 to 2075)
AUTHORS Finch,P.W., Aaronson,S.A. and Rubin,J.S.
TITLE Direct Submission
JOURNAL Submitted (09-JUN-1997) LCMB, DBS/NCI, 9000 Rockville Pike, Bethesda, MD 20892, USA
REMARK Sequence update by submitter
FEATURES
    Location/Qualifiers
        1..2075
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /chromosome="8"
            /map="8p11.1-p12"
            /clone="HSL"
            /dev_stage="embryonic"
            /tissue_type="lung"
            /cell_line="M426"
            /cell_type="fibroblast"
            303..1244
            /note="FRP"
            /codon_start=1
            /product="secreted frizzled-related protein"
            /db_xref="PID:g2213819"
            /translation="MGIGRSEGRGALGVLLAALVAGSASEYDVVSFQSDIGP
            YQSGREYTKPQCVDPADLRCHNVGKRVLPNLEHETMAEVKQOASWVPLLNK
            NCHAGQVFLCSLFAVCLDRPIYPCRWLCEAVRDSCEPVQFEGFYWEMLKCDKFP
            EDGVCIAMTPPNATEASKPGQTVCPDCELKSEAIIEHLCASEFALRMKIKVEKKE
            NGDKIIVPKKPKLGLPIKKDLKLVLYLKGADCPCHOLDNLSHHFLIMGRKRVKS
```

BASE COUNT	473 a	596 c	626 g	380 t	
ORIGIN	OYLLTAIHKWDKKNKEFKNFMKKMKNHCEPTFQSVFK"				
Query Match 5.1%; Score 103; DB 21; Length 2075;					
Best Local Similarity 73.5%; Pred. No. 1.77e-45;					
Matches 161; Conservative 0; Mismatches 58; Indels 0; Gaps 0;					
Db	449	CTTCTACCAAGCCACCTCAGTGCCTGGACATCCCGCGGAGCTGCGGCTGTGCCACAA	508		
Qy	351	CTTCTCTACAAGCGCAGCAATTGCAAGCCCACTCCGCGGCACTGTGCCACGG	410		
Db	509	CGTGGGTACAAGAAGATGTGTGTCGCCAACCTGTGTGGAGCAGACCATGGCGAGGT	568		
Qy	411	CATCGAATACGAGAACATGGCGTGCCTGCCAACCTGTGTGGCCACGAGACCATGAAGGAGT	470		
Db	569	GAACGACGCGCAGCAGCTGGGTGTCCTGCTCAACAGAACTGCCAGCGCGGACCCA	628		
Qy	471	GCTGGAGCGCGCGCGCTTGGATCCGCTGGTCATGAAGCAGTAGTCCCAACCGGACACAA	530		
Db	629	GGTCTCTCTGCTGCTGCTTTCGCGCGCGCTGTGCTGCTGA	667		
Qy	531	GAAGTTCTGTGCTGCTCTTGCGCCCGCTGCTGCTCGA	569		
RESULT	9				
LOCUS	AF056087	4469 bp	mRNA	PRI	08-APR-1998
DEFINITION	Homo sapiens secreted frizzled related protein mRNA, complete cds.				
ACCESSION	AF056087				
NID	93033550				
KEYWORDS	human.				
SOURCE	Homo sapiens				
ORGANISM	Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;				
REFERENCE	1 (bases 1 to 4469)				
AUTHORS	Zhou, Z. and Wang, J.				
TITLE	Upregulation of human secreted Frizzled homologue in apoptosis and its down regulation in breast tumors				
JOURNAL	Unpublished				
REFERENCE	2 (bases 1 to 4469)				
AUTHORS	Zhou, Z. and Wang, J.				
TITLE	Direct Submission				
JOURNAL	Submitted (29-MAR-1998) Radiohemmet Research Laboratory, CCK, Karolinska Institute, Solna, Sweden, Stockholm S-171 76, Sweden				
FEATURES	Location/Qualifiers				
source	1..4469				
	/organism="Homo sapiens"				
	/db_xref="taxon:9606"				
CDS	303...1244				
	/function="putative Wnt antagonist"				
	/note="FRP; frizzled homolog; cell cycle and apoptosis related; down regulated in breast tumorigenesis"				
	/codon_start=1				
	/product="secreted frizzled related protein"				
	/db_xref="PID:g3033551"				
	/translation="MGIGREGGRRGALGVLLAALLAVSGASEYDIVSFQSDIGL				
	YQSGRFYKPCQVDIPADUKLCNHYKKVILPNLLEHETMAEVKQKQASVPLLN				
	NCHAGTVFLCSLFAVPCVLDORPIYPCRWLCEAVRDCEPYMFFGFYPMELKCDKF				
	EGDVG IAMTPNATPEASKPQGTIVCPDNDLKESEAI IEHLCASEFALRMKIKVEYK				
	NGDKIVPKKPKLGLPIKIKKDLKLVLYLKNAGDCPCQHDNLNLSHHFLINGRKKV				
BASE COUNT	1121 a	1097 c	1120 g	1131 t	
ORIGIN	OYLLTAIHKWDKKNKEFKNFMKKMKNHCEPTFQSVFK"				
Query Match 5.1%; Score 103; DB 21; Length 4469;					
Best Local Similarity 73.5%; Pred. No. 1.77e-45;					
Matches 161; Conservative 0; Mismatches 58; Indels 0; Gaps 0;					
Db	449	CTTCTACCAAGCCACCTCAGTGCCTGGACATCCCGCGGAGCTGCGGCTGTGCCACAA	508		
Qy	351	CTTCTCTACAAGCGCAGCAATTGCAAGCCCACTCCGCGGCACTGTGCCACGG	410		

AUTHORS	TITLE
BURNER, J. J. and BARNER, J. J.	Recombinant fowlpox virus

BASE COUNT 51/ a 633 c 54/ g 3/8 c

BASE COUNT	517 a	633 c	547 g	378 t
	AYAWQKKNRNLHQAVRRWKHHRCPEQAGRKV*			

AUTHORS	Dorner, F., Scheiflinger, F. and Falkner, F. Gunter.
TITLE	Recombinant fowlpox virus

ORIGIN

Query Match 4.1%; Score 83; DB 16; Length 2075;
Best Local Similarity 71.1%; Pred. No. 2,548-32;
Matches 140; Conservative 0; Mismatches 57; Indels 0; Gaps 0;

Db 152 TGCACGCCATCCCGGCGAGCATGCCCTTGTCTATGACATCGTCTACTCGGAGATGAGG 211
QY 373 TGAAGCCCATCCCGGCCAACCTGCGAGTGTGCCACGGCATGAATACAGAACATCGCG 432

Db 212 ATCCCCAACTGCTGGAGCAGCAGACCATGCCGGAGGTGATCCAGCAGTCTCCAGCTGG 271
QY 433 CTGCCCCAACTGCTGGGCCACAGACCATGAAGAGAGGTGCTGGAGCAGCGCGGCTTGG 492

Db 272 CTGCCCCCTGCTGGCCAGGAGTGCATCCGACCGCAGGATTTCTCTGCTCCCTCTTC 331
QY 493 ATCCCGCTGGTCATGAAGCAGTGCACCCGGACACCAAGAGTCTCTGCTGCTCTTC 552

Db 332 GCGCCCATCTGCTCGA 348
QY 553 GCCCCCGTCTGCTCGA 569

RESULT 14
LOCUS HSU82169 2184 bp mRNA PRI 26-MAR-1997
DEFINITION Human frizzled homolog (FZD3) mRNA, complete cds.
ACCESSION U82169
NID g1906597
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 2184)
AUTHORS Wang, Y.-K., Samos, C.H., Peoples, R., Perez-Jurado, L.A., Nusse, R. and Francke, U.
TITLE A novel human homologue of the Drosophila frizzled wnt receptor gene binds wingless protein and is in the Williams syndrome deletion at 7q11.23
JOURNAL Hum. Mol. Genet. 6 (3), 465-472 (1997)
MEDLINE 97227293
REFERENCE 2 (bases 1 to 2184)
AUTHORS Wang, Y.-K., Peoples, R., Perez-Jurado, L.A. and Francke, U.
TITLE Direct Submission
JOURNAL Submitted (12-DEC-1996) Howard Hughes Medical Institute, Stanford Medical Center, Beckman Center B201, Stanford, CA 94305, USA
LOCATION/Qualifiers
FEATURES
source 1. .2184
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="7"
/map="7q11.23"
25. .1801
/gene="FZD3"
26. .1801
/gene="FZD3"
/note="similar to Drosophila frizzled; seven transmembrane receptor"
/codon_start=1
/product="frizzled homolog"
/db_xref="PID:g1906598"
/translation="MAVAPLRGALLLQWLAAGAALEIGRFPDGRGAAPQAVEI
PMCRGIGYNTMRPNLGHSTQGEAAAEAFAPLVQYCHSHLREFLCLSLVPMCTD
QVSTPIACRPMCEQALRCAPIMEQNFNGPDLSDCARLPTNDRPHALCMAPENAT
AGPAEPHKLGLMPAPRPPDGLPGAGSGTCENPEKFOYVEKSRSCAPRCGPV
VEVPSRRDKFLVWVMSALCFSTAPTLTLEPHERQYPERPIFILSMCVN
YSIAFLIRAVAGSQVACDQEGALVIEGLENTGCTLVLLIYFMSLSLWVYL
TLVLAAGKNGHEAIEAHGSYFMAAGLPALKTIVITLDRKVGADSLTGLCYVAS
TDAALTGFLVGLSYLGLSFLITGFVALPHIRKIMKTGTNTKEKLMKIGV
FSIYTPATCVICVYVERLNDFWRLRATQPCAAAGPGRDCSLPGGSVPTVA
VFMLKIFMSLVGITSVWVSSKTFQWLSYKRIAAAGRAKACRACRPGSVGRTH
CHYKAPTIVLHMTKTDPSLENPHL"

BASE COUNT 329 a 737 c 729 g 389 t

ORIGIN

Query Match 3.0%; Score 60; DB 21; Length 2184;
Best Local Similarity 59.8%; Pred. No. 7,848-18;
Matches 183; Conservative 0; Mismatches 123; Indels 0; Gaps 0;

Db 162 TGTCCCGGCGCATCGGTACAACTGACCGCATGCCCAACCTGCTGGGCGCACACGTCGC 221
QY 401 TGTCCCGGCGCATCGGTACAACTGACCGCATGCCCAACCTGCTGGGCGCACACGTCGC 460

Db 222 AGGCGGAGGGGCTGCGGAGTACGAGTTCGCGCGCTGGTGCAGTACGGTGCACACA 281
QY 461 TGAAGAGGTGCTGAGCAGCGCGGCTTGGATCCGCTGGTTCATGAAGCAGTGCACACC 520

Db 282 GCCACCTGCGCTTCTTCTGCTGCTGCTACGGGCCCATGTGTCACCGACCGAGTCTCGA 341
QY 521 CGGACACCAAGAGTTCCTGTGCTGCTTCGCGCCCGCTTCGCTCGATGACCTAGACG 580

Db 342 CGCCCATTCGCGCTGCGGCGCATGTGCGAGCAGGCGCGCTGCGCGCGCCCATCA 401
QY 581 AGCCATCCAGCGCATGCACTCGCTGCGTGCAGGTGAAGACCGCTGCGCGCGCGTCA 640

Db 402 TGGAGCAGTTCACTGCGGCTGCGCGGACTCGCTGACTGCGCGCGCTGCCACGCGCA 461
QY 641 TGTCCGCGCTTTCGCTTCCCTGCGCGGACATGCTTGAGTGCAGCGCTTCCCGCAGGACA 700

Db 462 AGCACC 467
QY 701 AGCACC 706

RESULT 15
LOCUS HSU43318 2334 bp mRNA PRI 24-FEB-1996
DEFINITION Human putative transmembrane receptor (frizzled 5) mRNA, complete cds.
ACCESSION U43318
NID g1151251
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 2334)
AUTHORS Wang, Y., Macke, J.P., Abella, B.S., Andreasson, K., Worley, P., Gilbert, D.J., Copeland, N.G., Jenkins, N.A. and Nathans, J.
TITLE A large family of putative transmembrane receptors homologous to the product of the Drosophila tissue polarity gene frizzled
J. Biol. Chem. 271 (8), 4468-4476 (1996)
JOURNAL 96224032
MEDLINE 2 (bases 1 to 2334)
REFERENCE Abella, B., Wang, Y., Macke, J.P. and Nathans, J.
AUTHORS Direct Submission
TITLE Submitted (14-DEC-1995) Jeremy Nathans, Molecular Biology and Genetics, Johns Hopkins Medical School, 725 N. Wolfe Street, Baltimore, MD 21205
LOCATION/Qualifiers
FEATURES
source 1. .2334
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="2"
/map="2q33-34"
/tissue_type="retina"
321. .2078
/gene="frizzled 5"
321. .2078
/gene="frizzled 5"
/note="putative transmembrane receptor"
/codon_start=1
/product="transmembrane receptor"
/db_xref="PID:g1151252"
/translation="MARPDSPAPPSLLLLLAQLVGRAAAASKAPVCOEITVPMCRGI
CYNLTHMPNQNHDTQDAGLEVHQFWLVEIQSPDLRFELCTNYTICILPDYHKPL
PFCRSVCRAGAGSPLMRQYGFAPWERNRSCDRLPVLGRDAEVLICMDYNRSEATTAPP

THIS PAGE BLANK (10/27)

W P S R E L

(TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPsrch_nn n.a. - n.a. database search, using Smith-Waterman algorithm
Run on: Wed Aug 5 07:34:48 1998; MasPar time 322.83 Seconds
Tabular output not generated.

Title: >US-08-848-439-1
Description: (1-2027) from US08848439.seq
Perfect Score: 2027
N.A. Sequence: 1 GAATTCGCGCTTCATGGCT.....AAAAAAGGCGCGC 2027
Comp: CTTAAGCGCGAAGTACCGGA.....TTTTTTTTTCCCGCGC

Scoring table: TABLE default
Gap 6

Nmatch STD : Dbase 0; Query 0

Searched: 188442 seqs, 58026449 bases x 2

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: n-geneseq32
1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7
8:part8 9:part9 10:part10 11:part11 12:part12 13:part13
14:part14 15:part15 16:part16 17:part17 18:part18
19:part19 20:part20 21:part21 22:part22 23:part23
24:part24 25:part25 26:part26 27:part27 28:part28
29:part29 30:part30 31:part31 32:part32 33:part33
34:part34 35:part35 36:part36 37:part37 38:part38
39:part39 40:part40

Statistics: Mean 9.846; Variance 6.883; scale 1.431

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result		Query		Match		Length		DB		ID		Description		Pred. No.	
1	201	9.9	204	20	T23535	Human gene signature	1.01e-97					Human gene signature	1.01e-97		
2	58	2.9	2334	37	T89889	Human frizzled gene 5	6.84e-16					Human frizzled gene 5	6.84e-16		
3	56	2.8	2421	37	T89892	Mouse frizzled gene 8	7.60e-15					Mouse frizzled gene 8	7.60e-15		
4	44	2.2	91	9	Q51746	Oligonucleotide probe	9.73e-09					Oligonucleotide probe	9.73e-09		
5	45	2.2	172	32	T76363	Human interleukin 8 a	3.11e-09					Human interleukin 8 a	3.11e-09		
6	43	2.1	204	1	N81164	Base substituted E.co	2.88e-08					Base substituted E.co	2.88e-08		
7	41	2.0	91	9	Q51746	Oligonucleotide probe	8.76e-07					Oligonucleotide probe	8.76e-07		
8	40	2.0	100	32	T76186	Human IL4 receptor an	8.76e-07					Human IL4 receptor an	8.76e-07		
9	40	2.0	114	12	Q70466	Generic DNA sequence	8.76e-07					Generic DNA sequence	8.76e-07		
10	40	2.0	114	12	Q70468	Generic DNA sequence	8.76e-07					Generic DNA sequence	8.76e-07		
11	40	2.0	204	1	N81164	Base substituted E.co	8.76e-07					Base substituted E.co	8.76e-07		
12	39	1.9	114	12	Q70469	Generic DNA sequence	2.65e-06					Generic DNA sequence	2.65e-06		
13	38	1.9	114	12	Q70467	Generic DNA sequence	7.95e-06					Generic DNA sequence	7.95e-06		

14	38	1.9	114	12	Q70465	Generic DNA sequence	7.95e-06
15	39	1.9	114	12	Q70469	Generic DNA sequence	2.65e-06
16	38	1.9	114	12	Q70468	Generic DNA sequence	7.95e-06
17	38	1.9	114	12	Q70466	Generic DNA sequence	7.95e-06
18	39	1.9	178	32	T76405	Human endothelin-1 an	2.65e-06
19	39	1.9	190	32	T76452	Chymase antisenese oil	2.65e-06
20	36	1.8	114	12	Q70470	Generic DNA sequence	6.96e-05
21	36	1.8	114	12	Q70472	Generic DNA sequence	6.96e-05
22	36	1.8	114	12	Q70473	Generic DNA sequence	6.96e-05
23	34	1.7	114	12	Q70471	Generic DNA sequence	5.87e-04
24	34	1.7	114	12	Q70465	Generic DNA sequence	5.87e-04
25	34	1.7	114	12	Q70471	Generic DNA sequence	5.87e-04
26	34	1.7	114	12	Q70470	Generic DNA sequence	5.87e-04
27	34	1.7	114	12	Q70471	Generic DNA sequence	5.87e-04
28	34	1.7	250	32	T76438	Substance P antisenese	5.87e-04
29	35	1.7	317	32	T76274	Human neutrophil elas	2.03e-04
30	32	1.6	114	12	Q70472	Generic DNA sequence	4.71e-03
31	32	1.6	178	32	T76405	Human endothelin-1 an	4.71e-03
32	32	1.6	264	32	T76445	Substance P receptor an	4.71e-03
33	30	1.5	88	32	T76170	Human IL3 receptor an	3.65e-02
34	31	1.5	168	32	T76270	Human MDNCF antisenese	1.32e-02
35	31	1.5	190	32	T76452	Chymase antisenese oil	1.32e-02
36	30	1.5	317	32	T76274	Human neutrophil elas	3.65e-02
37	30	1.5	565	6	Q35072	RCV envelope region n	3.65e-02
38	30	1.5	3871	2	N71302	HSV-1 gB and surround	3.65e-02
39	29	1.4	70	32	T76184	Human IL4 antisenese o	9.94e-02
40	29	1.4	128	32	T76233	Human IL6 antisenese o	9.94e-02
41	28	1.4	162	32	T76307	Human RANTES antisenese	2.67e-01
42	29	1.4	172	32	T76363	Human interleukin 8 a	9.94e-02
43	28	1.4	1466	2	N91353	Vascular anti-coagula	2.67e-01
44	28	1.4	1840	11	Q56611	Feline zona pellucida	9.94e-02
45	29	1.4	2259	37	T89891	Mouse frizzled gene 7	9.94e-02

ALIGNMENTS

RESULT 1
ID T23535 standard; cDNA to mRNA; 204 BP.
AC T23535
DT 02-SEP-1996 (first entry)
DE Human gene signature HUMGS05382.
KW Gene signature; messenger RNA; mRNA; relative abundance; frequency;
KW human; cloning; mapping; non-biased library; diagnosis; detection;
KW cell typing; abnormal cell function; ss.
OS Homo sapiens.
PN W09514772-A1.
PD 01-JUN-1995.
PF 11-NOV-1994; J01916.
PR 12-NOV-1993; JP-355504.
PA (MATS/) MATSUBARA K.
PA (OKUB/) OKUBO K.
PI Matsubara K. Okubo K.
DR WPI; 95-206931/27.
PT Identifying gene signatures in 3'-directed human cDNA library - e.g.
PT for diagnosis of abnormal cell function, by preparing cDNA that
PT reflects relative abundance of corresp. mRNA in specific human
PT tissues
PS Claim 1; Page 1393; 2245pp; Japanese.
CC A single-stranded DNA (or its complementary strand or the corresp.
CC double-stranded DNA) which comprises one of the 7837 "GS" sequences
CC given in T19001-T26837 and which is able to hybridise to part of
CC human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
CC sequences were obtained from 3'-directed cDNA libraries prepared
CC from various human tissues; synthesis of cDNA was initiated from the
CC 3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
CC untranslated sequence is unique to a particular mRNA species, almost
CC all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
CC is constructed so as to reflect accurately the relative abundance of
CC different mRNAs in the particular tissue from which it was derived.
CC The appearance frequency of a given GS in a cDNA library can be
CC determined (esp. using primers and probes derived from the GS
CC sequences) as a means of diagnosing abnormal cell function or for
CC recognising different cell types.

SQ Sequence 204 BP; 54 A; 59 C; 35 G; 55 T;

Query Match 9.9%; Score 201; DB 20; Length 204;
 Best Local Similarity 99.5%; Pred. No. 1.01e-97;
 Matches 201; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

D 1 gatctcagctccgcttcccaagcacactcctagctcagctcagctcagctggcagct 60
 |||||
 QY 1207 GATCTCAGCTCCGCTTCCCAAGCACACTCTAGCTGCTCGAGTCTCAGCTTGGGCACT 1266
 |||||

D 61 tccccctgccttttgacgtttgcatcccccagcatttccctgagttataaaggccacaggag 120
 |||||
 QY 1267 TCCCCCTGCCCTTTTGACAGTTTGATCCCCAGCATTTCTTCTGAGTTATAAGGCCACAGGAG 1326
 |||||

D 121 tggatacgtgttttcacctaagaagaaagccaccacccaattctgtagaaatattcaaac 180
 |||||
 QY 1327 TGGATAGCTGTTTTCACCTTAAGGAAAGAGCCACCCGAACTTTGTAGAAATATTCAAAC 1386
 |||||

D 181 antaaatcatgaattttta 202
 |||||

QY 1387 AATAAATCATGAATATTITA 1408
 |||||

RESULT 2

ID T89889 standard; DNA; 2334 BP.
 AC T89889;
 DT 27-APR-1998 (first entry)
 DE Human frizzled gene 5 (Hfz5) encoding a Wnt receptor.
 KW Wnt receptor; human frizzled gene 5; Hfz5 gene;
 KW signal transduction; cancer; cell growth; cell proliferation;
 KW mammary tumour; oncogene; therapy; ds.
 OS Homo sapiens.
 FH Key Location/Qualifiers
 FT CDS 321..2078
 FT /*tag= a
 PN WO9739357-A1.
 PD 23-OCT-1997.
 PF 11-APR-1997; U06049.
 PR 12-APR-1996; US-015307.
 PA (UYJO) UNIV JOHNS HOPKINS.
 PA (STRD) UNIV LELAND STANFORD JUNIOR.
 PI Andrew D, Bhanot P, Brink M, Hsieh J, Nathans J,
 PI Nusse R, Samos CH, Wangy;
 DR WPI; 97-526631/48.
 DR P-PSDB; W31271.
 PT Identification of Wnt receptor binding modulators - useful for
 PT treatment of cancer and growth, development or proliferation related
 PT disorders
 PS Disclosure; Page 36-37; 61pp; English.
 CC This sequence comprises novel human frizzled gene 5 (Hfz5)
 CC that encodes a transmembrane receptor, frizzled-5 (see W31771), a
 CC Wnt receptor (WntR). Novel frizzled family members have been
 CC identified in Drosophila, mouse, human and Caenorhabditis (see
 CC T89885-92) and are considered to be Wnt receptors. Wnt receptors
 CC can be used in a novel, claimed method of screening for compounds
 CC which modulate the binding of a Wnt polypeptide (secreted proteins
 CC involved in cell-to-cell signalling) to a Wnt receptor. Wnt is
 CC involved in (mammary) cancer and other processes involving growth,
 CC development and proliferation (both normal and abnormal).
 CC Modulators identified by the claimed method are useful for
 CC treatment of diseases related to these conditions.
 SQ Sequence 2334 BP; 356 A; 803 C; 736 G; 439 T;

Query Match 2.9%; Score 58; DB 37; Length 2334;
 Best Local Similarity 59.9%; Pred. No. 6.84e-16;
 Matches 175; Conservative 0; Mismatches 117; Indels 0; Gaps 0;

D 439 tgtgcgcggcagctcggtacacactgacgcacatgcccaaccagttcaaacacgacacgc 498
 |||||
 QY 401 TGTGCCACGGCATCGAATACAGAAATGCGGCTGCCCAACCTGCTGGGCCACGAGACCA 460
 |||||

D 499 agcacaggcgccctggaggtgcacagttctggtgcgcgtggtggagatccaatgtcgc 558
 |||||

QY 461 TGAAGGAGGTGCTGGAGCAGGCGCGCTTGGATCCCGTGTGATGAAGCAGTCCACC 520
 |||||

D 559 cggacctgcgcttcttcctactatgcactatgtacacgcccactgtctgtccgcactaccaca 618
 |||||

QY 521 CGGACACCAAGAAGTTCCCTGCTGCTCTTGGCCCCCGCTGCTGCTCGATGACCTAGACG 580
 |||||

D 619 agcgcgtgcgcgcctcgcgctggtgtgcagcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 678
 |||||

QY 581 AGACCATCCAGCCATGCCACTGCTCTGCTGAGGTGAAGGACCGCTGCGCCCCGGTCA 640
 |||||

D 679 tgcgcagctacggtcttcgctggtccgcgagcgcagctgagctgcgacgcgcctccc 730
 |||||

QY 641 TGTCCGCCCTTGGCTTCCCTTGGCCCGACATGCTTGTGAGTGGACCGCTTCCC 692
 |||||

RESULT 3

ID T89892 standard; DNA; 2421 BP.
 AC T89892;
 DT 27-APR-1998 (first entry)
 DE Mouse frizzled gene 8 (Mfz8) encoding a Wnt receptor.
 KW Wnt receptor; mouse frizzled gene 8; Mfz8 gene;
 KW signal transduction; cancer; cell growth; cell proliferation;
 KW mammary tumour; oncogene; therapy; ds.
 OS Mus musculus.
 FH Key Location/Qualifiers
 FT CDS 188..2245
 FT /*tag= a
 PN WO9739357-A1.
 PD 23-OCT-1997.
 PF 11-APR-1997; U06049.
 PR 12-APR-1996; US-015307.
 PA (UYJO) UNIV JOHNS HOPKINS.
 PA (STRD) UNIV LELAND STANFORD JUNIOR.
 PI Andrew D, Bhanot P, Brink M, Hsieh J, Nathans J,
 PI Nusse R, Samos CH, Wangy;
 DR WPI; 97-526631/48.
 DR P-PSDB; W31274.
 PT Identification of Wnt receptor binding modulators - useful for
 PT treatment of cancer and growth, development or proliferation related
 PT disorders
 PS Disclosure; Page 46-47; 61pp; English.
 CC This sequence comprises novel mouse frizzled gene 8 (Mfz8)
 CC that encodes a transmembrane receptor, frizzled-8 (see W31274),
 CC a Wnt receptor (WntR). Novel frizzled family members have been
 CC identified in Drosophila, mouse, human and Caenorhabditis (see
 CC T89885-92) and are considered to be Wnt receptors. Wnt receptors
 CC can be used in a novel, claimed method of screening for compounds
 CC which modulate the binding of a Wnt polypeptide (secreted proteins
 CC involved in cell-to-cell signalling) to a Wnt receptor. Wnt is
 CC involved in (mammary) cancer and other processes involving growth,
 CC development and proliferation (both normal and abnormal).
 CC Modulators identified by the claimed method are useful for
 CC treatment of diseases related to these conditions.
 SQ Sequence 2421 BP; 403 A; 772 C; 779 G; 467 T;

Query Match 2.8%; Score 56; DB 37; Length 2421;
 Best Local Similarity 59.4%; Pred. No. 7.60e-15;
 Matches 177; Conservative 0; Mismatches 121; Indels 0; Gaps 0;

D 306 tgcgcttgcgaaggcctcggttcaactacactacactacacacacacacacacacacac 365
 |||||

QY 395 TCCAGCTGTGCCACGGCATCGAATACCAAGACATCGCGCTGCCAACCTGTGGGCCACG 454
 |||||

D 366 acacgcaagatgagcgcgccctagaggtgcacagttttgcccgtgtggagatacagt 425
 |||||

QY 455 AGACCATGAAGAGGTGCTGGAGCAGCGCGGCTTGGATCCCGTGTGATGAAGCAGT 514
 |||||

D 426 gctccccggacctcaagtcttcttctgttagatgtacacgccatctgcctggaggact 485
 |||||

QY 515 GCCACCCGGACACCAAGAAGTTCCTGCTGCTGCTTCCGCCCGCTGCTCGATGACC 574
 |||||

D 486 acaagaagccttcgcgcctgtgctgtgtgtgaacgcgcgcgcgcgcgcgcgcgcgcgcgc 545
 |||||

RESULT 5
ID T76363 standard; DNA: 172 BP.
AC 176363;
DE Human interleukin 8 antisense oligonucleotide.
DE 15-SEP-1997 (first entry)
DE DE Human interleukin 8 antisense oligonucleotide.
DE DE Human interleukin 8 antisense oligonucleotide.
KW Asthma; airway epithelium; adenosine free; cystic fibrosis;
KW chronic obstructive pulmonary disease; bronchitis; ss.
KW Synthetic.
PN WO9640162-A1.
PD 19-DEC-1996.
PF 06-JUN-1996; U09306.
PR 07-JUN-1995; US-474497.
PA (UYEC-) UNIV EAST CAROLINA.
PI Metzger WJ, Nyce JW;
DR MPI: 97-05187/05.
PT Treatment of airway diseases such as asthma - by topically applying
PT adenosine-free antisense oligo:nucleotide to airway epithelium of
PT subject
PS Claim 5: Page 36: 71pp: English.
CC A method for treating airway disease in a subject has been produced.
CC which involves the topical administration of an essentially adenosine
CC free antisense oligonucleotide (ON) to the airway epithelium of the
CC subject. The present sequence is an antisense oligonucleotide specific
CC for the human interleukin 8, targeted at the initiation codon. The
CC method can be used to treat airway diseases such as cystic fibrosis,
CC asthma, chronic obstructive pulmonary disease, bronchitis and other
CC airway diseases characterised by an inflammatory response. By
CC eliminating adenosine from the antisense ON, its liberation upon
CC antisense degradation is prevented, thereby preventing adenosine-
CC induced bronchoconstriction in patients with hyper-reactive airways.
SQ Sequence 172 BP: 0 A: 35 C: 42 G: 39 T;

RESULT	7
ID	Q51746 standard; cDNA; 91 BP.
AC	Q51746;
DT	31-MAY-1994 (first entry)

RESULT	10
ID	Q70468
AC	Q70468
DT	05-APR


```
DE Generic DNA sequence to generate a random TSAR peptide library.
KW TSAR; totally synthetic affinity reagent; synthetic; binding domain;
KW effector domain; concatenated heterofunctional protein; linker;
KW direct; rapid; detection; screening; treatment; generic; ss.
OS Synthetic.
FH Key Location/Qualifiers
FT misc_feature 55..60
FT /*tag= a
FT /*note= "this sequence represents 'z'; z can be a
FT sequence of 6, 9 or 12 nucleotides (see
FT comments)"
PN WO9418318-A.
PD 18-AUG-1994.
PF 01-FEB-1994. U00977.
PR 01-FEB-1993; US-013416.
PR 30-DEC-1993; US-176500.
PR 31-JAN-1994; US-189331.
PA (UUNC-) UNIV NORTH CAROLINA.
PI Fowlkes DM, Kay BK;
WPI: 94-279739/34.
DR P-PSDB: R65154.
DR Identifying proteins or peptide(s) which bind a ligand - by
PT screening a recombinant vector library expressing fusion proteins
PT comprising a binding domain and an effector domain
PS Disclosure: Page 35; 255pp; English.
CC Q70468 is a generic DNA sequence used to generate random TSAR (Totally
CC Synthetic Affinity Reagents) peptides. This generic formula can also be
CC represented as follows: X(NNB)11(TGC)(NNB)8Z(NNB)7(TGC)(NNB)10Y. X
CC and Y are flanking restriction sites (X is not the same as Y) that are
CC not specified further. Other generic sequences are shown in Q70466-68.
CC Other specific peptides generated by these generic sequences are shown in
CC Q70469-74. TSARs are concatenated heterofunctional proteins or peptides,
CC comprising at least two functional regions - a binding domain with
CC affinity for a ligand and a second effector peptide portion that is
CC chemically or biologically active. They may further comprise a linker
CC peptide between the 2 domains. The oligonucleotides are also designed so
CC that the expressed peptide contains 2 or 4 cysteine residues positioned
CC in, or flanking, the unpredicted or variant residues. These residues
CC confer some degree of conformational rigidity to the peptides. The TSARs
CC or compns. comprising a TSAR binding domain can be used in vivo to
CC deliver a chemically or biologically active moiety, eg. metal ion,
CC radioisotope, peptide, toxin or enzyme, to the specific target or on the
CC cell. They can also replace the function of macromolecules, eg.
CC monoclonal or polyclonal antibodies and therefore circumvent the need
CC for complex methods of hybridoma formation or in vivo antibody
CC production. The TSARs are easily characterised and have designed activity
CC allowing direct and rapid detection in a screening process.
SQ Sequence 114 BP; 0 A; 2 C; 2 G; 2 T;

Query Match 2.0%; Score 40; DB 12; Length 114;
Best Local Similarity 5.4%; Pred. No. 8.76e-07;
Matches 6; Conservative 34; Mismatches 72; Indels 0; Gaps 0;

Db 3 bnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnnn 62
QY 199 CCGCCTCGCCCTCCCGGCTCGCTCCCTCTCGCCCTCGGCGCGCGCCACGATG 258
Db 63 bnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnn 114
QY 259 CTCAGGGCCCTGGCTCGCTGCTGCTCTCTCTCGCTCGCCTCGCCTCGCC 310

RESULT 11
ID N81164 standard; DNA; 204 BP.
AC N81164;
DT 08-NOV-1990 (first entry)
DE Base substituted E.coli beta-galactosidase alpha-fragment.
KW E.coli beta galactosidase alpha-fragment; base substitutions; ss.
OS Escherichia coli.
FH Key Location/Qualifiers
FT misc_feature 19..69
FT /*tag= a
FT /function= multiple cloning site
```

```
FT primer_bind 187..204
FT /*tag= b
PN EP-285123-A.
PD 05-MAY-1988.
PF 30-MAR-1988; 105163.
PR 03-APR-1987; US-034819.
PA (SUSO) SUOMEN SOKERI OY.
PI Lehtovaara P, Knowles J, Koivula A, Bamford J, Reinikainen T;
WPI: 88-279927/40.
DR Introducing random point mutations into nucleic acids -
PT by prepn of single stranded template, annealing a primer, elongation,
PT misincorporation, completion of molecules and screening.
PS Disclosure: P; English.
CC Random point mutations were introduced into the alpha fragment of
CC E.coli beta-galactosidase. The wild type sequence was obtained as a
CC single stranded template and an oligonucleotide was hybridised to
CC it to generate a popn of DNA molecules which terminate at all
CC possible nucleotide positions within a specified region. The
CC variable 3' ends generated in this way are used as primers for
CC reverse transcriptase. Nucleotides are misincorporated by the
CC transcriptase and the molecules are completed to forms that can be
CC amplified and then expressed in a suitable host-vector system.
CC The sequence covers all 176 diff base substitutions, most of which
CC occurred singularly in any given mutant.
CC See also F80375.
SQ Sequence 204 BP; 21 A; 47 C; 17 G; 11 T; 108 Others;

Query Match 2.0%; Score 40; DB 1; Length 204;
Best Local Similarity 15.6%; Pred. No. 8.76e-07;
Matches 20; Conservative 57; Mismatches 51; Indels 0; Gaps 0;

Db 59 ggyywcgagcygcaaycdhvcgcygrrtthhyrrmbnvyrdynrsdaawccyr 118
Cp 343 GCCCAAGAGGAAGAGCGCGCGCGAGCCAGCAGTCGCGAGGCGGAGAGCA 284
Db 119 rsvkydcynachdhvyybbvynvnhnncnccbnhvnbnhnnhrnwayvrdh 178
Cp 283 GCAGGAGGAGGAGGCGCGCGCGCGAGCAGTCGTCGCGCGGCGGCGGCGGAGGA 224
Db 179 arddgvhc 186
Cp 223 GCGGAGCC 216

RESULT 12
ID Q70469 standard; DNA; 114 BP.
AC Q70469;
DT 07-APR-1995 (first entry)
DE Generic DNA sequence to generate a random TSAR peptide library.
KW TSAR; totally synthetic affinity reagent; synthetic; binding domain;
KW effector domain; concatenated heterofunctional protein; linker;
KW direct; rapid; detection; screening; treatment; generic; ss.
OS Synthetic.
FH Key Location/Qualifiers
FT misc_feature 55..60
FT /*tag= a
FT /*note= "this sequence represents 'z'; z can be a
FT sequence of 6,9 or 12 nucleotides (see
FT comments)"
PN WO9418318-A.
PD 18-AUG-1994.
PF 01-FEB-1994. U00977.
PR 01-FEB-1993; US-013416.
PR 30-DEC-1993; US-176500.
PR 31-JAN-1994; US-189331.
PA (UUNC-) UNIV NORTH CAROLINA.
PI Fowlkes DM, Kay BK;
WPI: 94-279739/34.
DR Identifying proteins or peptide(s) which bind a ligand - by
PT screening a recombinant vector library expressing fusion proteins
PT comprising a binding domain and an effector domain
PS Disclosure: Page 35; 255pp; English.
CC Q70469 is a generic DNA sequence used to generate random TSAR peptide
```


THIS PAGE BLANK (USPTO)

WATER

(TW)

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPsrch_nn n.a. - n.a. database search, using Smith-Waterman algorithm
Run on: Wed Aug 5 06:20:04 1998; MasPar time 2338.81 Seconds
1213.820 Million cell updates/sec
Tabular output not generated.

Title: >US-08-848-439-1
Description: (1-2027) from US08848439.seq
Perfect Score: 2027
N.A. Sequence: 1 GAATTCGGCCTTCATGGCT.....AAAAAAGGCGCGC 2027
Comp: CTTAAGCGGAGTACCGGA.....TTTTTTTTTCGCCGGC

Scoring table:
Gap 6
TABLE default

Nmatch STD : Dbase 0: Query 0

Searched: 1832099 segs, 700269816 bases x 2

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: emb1-est55
1:em_est10 2:em_est11

Database: genbank-est107
3:gb_est1 4:gb_est10 5:gb_est11 6:gb_est12 7:gb_est13
8:gb_est14 9:gb_est15 10:gb_est16 11:gb_est17
12:gb_est18 13:gb_est19 14:gb_est2 15:gb_est20
16:gb_est21 17:gb_est22 18:gb_est23 19:gb_est3
20:gb_est4 21:gb_est5 22:gb_est6 23:gb_est7 24:gb_est8
25:gb_est9 26:gb_gss

Statistics: Mean 12.269; Variance 3.137; scale 3.912

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	553	27.3	579	AA528226	nh92a05.s1 NCI_CGAP.Br	0.00e+00
2	479	23.6	505	24	ze88f06.s1 Soares feta	0.00e+00
3	473	23.3	480	21	Yy64d09.r1 Homo sapien	0.00e+00
4	472	23.3	505	23	ze03g10.r1 Soares feta	0.00e+00
5	468	23.1	473	8	nh88g03.s1 NCI_CGAP.Br	0.00e+00
6	468	23.1	474	15	of62a09.s1 NCI_CGAP.Co	0.00e+00
7	453	22.3	460	23	zd70e11.s1 Soares feta	0.00e+00
8	434	21.4	443	7	zx08c06.s1 Soares tota	0.00e+00
9	433	21.4	482	18	ch10h11.s1 NCI_CGAP.Co	0.00e+00
10	425	21.0	427	5	zw51g08.s1 Soares tota	0.00e+00
11	420	20.7	518	17	on47d08.s1 NCI_CGAP.Co	0.00e+00
12	416	20.5	446	21	Yy64d09.s1 Homo sapien	0.00e+00
13	413	20.4	462	23	zd70e11.r1 Soares feta	0.00e+00

C	14	409	20.2	409	7	AA449749	zx07e10.s1 Soares tota	0.00e+00
C	15	408	20.1	445	20	N32415	Yw83c06.s1 Homo sapien	0.00e+00
C	16	407	20.0	562	7	AA449032	zx07e10.r1 Soares tota	0.00e+00
C	17	405	20.0	405	8	AA505877	ni01a05.s1 NCI_CGAP.Br	0.00e+00
C	18	402	19.8	408	18	AA976403	Qq8f09.s1 NCI_CGAP_Ki	0.00e+00
C	19	388	19.1	388	7	AA528219	nj16h09.s1 NCI_CGAP_Pr	0.00e+00
C	20	366	18.1	422	5	AA431858	zw51g08.r1 Soares tota	0.00e+00
C	21	366	18.1	458	24	AA105749	ml84a03.r1 Stratagene	0.00e+00
C	22	362	17.9	378	10	AA429960	zw58g10.r1 Soares tota	0.00e+00
C	23	363	17.9	433	22	AA024772	ze76g10.s1 Soares feta	0.00e+00
C	24	359	17.7	426	20	H87071	Ys74d07.r1 Homo sapien	0.00e+00
C	25	350	17.3	359	4	AA295018	EST100474 Pancreas tum	0.00e+00
C	26	337	16.6	406	19	H16121	Y128c05.r1 Homo sapien	0.00e+00
C	27	331	16.3	340	6	HUM296B07B	Human aorta cDNA 5'-en	0.00e+00
C	28	329	16.2	396	25	AA073862	mm97f07.r1 Stratagene	0.00e+00
C	29	328	16.2	414	22	AA024771	ze76g10.r1 Soares feta	0.00e+00
C	30	323	15.9	329	12	AA705737	zf41b02.s1 Soares feta	0.00e+00
C	31	322	15.9	434	19	H44092	YO73a07.r1 Homo sapien	0.00e+00
C	32	321	15.8	321	4	AA335712	EST40158 Epididymus Ho	0.00e+00
C	33	320	15.8	328	4	AA295688	EST100887 Pancreas tum	0.00e+00
C	34	321	15.8	469	19	H15818	Y128c05.s1 Homo sapien	0.00e+00
C	35	319	15.7	326	20	N32424	Yw83e07.s1 Homo sapien	0.00e+00
C	36	316	15.6	330	6	HUM230F10B	Human aorta cDNA 5'-en	0.00e+00
C	37	313	15.4	344	21	N56835	Yw83c06.r1 Homo sapien	0.00e+00
C	38	299	14.8	301	6	HUM303B01B	Human aorta cDNA 5'-en	0.00e+00
C	39	301	14.8	301	10	AA602994	np31g01.s1 NCI_CGAP_Pr	0.00e+00
C	40	299	14.8	313	6	HUM304B12B	Human aorta cDNA 5'-en	0.00e+00
C	41	298	14.7	303	6	HUM347H05B	Human aorta cDNA 5'-en	0.00e+00
C	42	296	14.6	300	4	AA347786	EST54420 Fetal heart I	0.00e+00
C	43	294	14.5	300	6	HUM307G04B	Human aorta cDNA 5'-en	0.00e+00
C	44	292	14.4	302	4	AA330758	EST34493 Embryo, 6 wee	0.00e+00
C	45	290	14.3	293	6	HUM238C04B	Human aorta cDNA 5'-en	0.00e+00

ALIGNMENTS

RESULT 1
LOCUS AA528226 579 bp mRNA EST 05-AUG-1997
DEFINITION nh92a05.s1 NCI_CGAP_Br1.1 Homo sapiens cDNA clone IMAGE:965936,
mRNA sequence.
ACCESSION AA528226
NID 92270295
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryote; mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homnidae;
Homo.
REFERENCE 1 (bases 1 to 579)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert_Strausberg@nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D.
DNA Sequencing by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.lnl.gov/bbrp/image/image.html
Insert Length: 1268 Std Error: 0.00
Seq primer: -40m13 fwd. ET from Amersham
High quality sequence stop: 206.
Location/Qualifiers
1. .579
/organism="Homo sapiens"
/note="vector: pT73D-Pac (Pharmacia) with a modified

RESULT	2	EST	01-FEB-1997
LOCUS	AA071459	505 bp	mrna
DEFINITION	ze89f06.sl Soares fetal heart		
	366083 3',		mrna sequence.
ACCESSION	AA071459		
NID	91578900		
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
	Eukaryotae; mitochondrial		eukaryotes; Metazoa; Chordata;

[illegible]

Qy	1636	TGAGACTCAGTGCTAAGTCTTTACACAAATTCATCTTTTATACCTTCAATGGGA	CTTA	1695
Db	301	AACTGTTACATGTATCACATTC	CCAGCTACAATPACTTC	CACTTTATTAGAACACATAACC
Qy	1696	AACTGTTACATGTATCACATTC	CAGCTACAATPACTTC	CACTTTATTAGAACACATAACC
Db	361	ATTTCATATGACATGATTTCTT	CAAGCTAAAAGCGAAACATATAA	TTTTATATAATTTGACCT
Qy	1756	ATTTCTATGACATGATTTCTT	CAAGTAAAGCGAAACATATAA	TTTTATATAATTTGACCT
Db	421	GAGTACTTTAAGCCTTGTTTAA	ACATTTCTACTTAACCTTTTG	CAAAATTAACCCCAT
Qy	1816	GAGTACTTTAAGCCTTG-TTTA	AAACATTTCTACTTAACCTTT	TGCAAAATTAACCCCAT
RESULT	4			
LOCUS	w92531	505 bp	mRNA	EST
DEFINITION	ze03g10.r1	Soares fetal heart	NbHH19w	Homo sapiens CDNA clone

	357954, 5', mRNA sequence.
ACCESSION	W92531
NID	g1425099
KEYWORDS	EST.
SOURCE	human.
ORGANISM	Homo sapiens
	Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
	Vertebrata; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE	1 (bases 1 to 505)
AUTHORS	Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Travaskis,E., Waston,R., Williamson,A., Workalembo,P. and

TITLE	COMMENT
The WashU-Merck EST Project Unpublished (1995)	Contact: Wilson RK WashU-Merck EST Project Washington University School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108 Tel: 314 286 1800 Fax: 314 286 1810 Email: est@watson.wustl.edu This clone is available royalty-free through LLNL : contact the

FEATURES
source
Location/Qualifiers
1. .505
High quality sequence stop: 389.
Seq primer: mob.REGA-E1
Insert length: 692 Std Error: 0.00
IMAGE Consortium (info@image.lnl.gov) for further information.

modified polylinker: Site 1: Not I; Site 2: Eco RI; 1 strand cDNA was primed with a Not I - oligo(dT) primer: TGTTCACATCTGAGTGCGGGCCGCATCTTTTTTTTTTTT 3' |

```

normalization to a Cot = 5. Library constructed by
M. Ratina Ronaldo. This library was constructed from the
same fetus as the fetal lung library, Soares fetal lung
NbHL19w."
/db_xref="taxon:9606"
/clone="357954"
/clone_lib="Soares fetal heart NbHL19w"
/sex="unknown"
/dev_stage="19 weeks"
/lab_host="DH10B (ampicillin resistant)"

```

mRNA	BASE COUNT	ORIGIN
<1. .>505	149 a	95 c
	81 g	180 t

Query Match	23.3%	Score 472;	DB 23;	Length 505;
Best Local Similarity	98.8%	Pred. No. 0.00e+00;		
Matches	498;	Conservative 0;	Mismatches 2;	Indels 4;
Gaps				
Db	1	ATTTTATGAAGTTTAAAAATACGTCACATTAAAGCTAGTTTGAATAGTGCAACTG 60		
QY	1402	ATTTTATGAAGTTTAAAAATACGTCACATTAAAGCTAGTTTGAATAGTGCAACTG 1461		
Db	61	ACTTGGCTCTGGTTGGTTGTTTGTGTTTGTAGTCAGCTGATTTTCACTTCCCACGTG 120		
QY	1462	ACTTGGCTCTGGTTGGTTGTTTGTGTTTGTAGTCAGCTGATTTTCACTTCCCACGTG 1521		
Db	121	AGGTGTGCATAAGCATGCAAAATTCGTTCAAATTTCTCTGTGGGCCCAACTTGTGGGTACA 180		
QY	1522	AGGTGTGCATAAGCATGCAAAATTCGTTCAAATTTCTCTGTGGGCCCAACTTGTGGGTACA 1581		
Db	181	AACCCCTGTTGAGATAAGCTGGCTGTATCTCAACATCTTCATCAGCTCCAGACTGAGAC 240		
QY	1582	AACCCCTGTTGAGATAAGCTGGCTGTATCTCAACATCTTCATCAGCTCCAGACTGAGAC 1641		
Db	241	TCAGTGTCTAAGTCTTTACAACAATTCATCATTTTTATACCTTCAATGGGAACCTTAAACTGT 300		
QY	1642	TCAGTGTCTAAGTCTTACACAAATTCATCATTTTTATACCTTCAATGGGAACCTTAAACTGT 1701		
Db	301	TACATGTATCACATTCGAGTCAAGTCAACTACTTCCTCATTTTAAAGAAGCAATTAACCATTT 360		
QY	1702	TACATGTATCACATTCGAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGT 1759		
Db	361	CTATAGCATGATTTCTTCAAGTAAAGGCAAGGCAAGATATAAATTTTAAATGACTTGAGT 420		
QY	1760	CTATAGCATGATTTCTTCAAGTAAAGGCAAGGCAAGATATAAATTTTAAATGACTTGAGT 1819		
Db	421	ACTTTAAGCCCTGTTTAAAAACATTTCTTACTTAACTTTTGCAAAATTAACCCCATTTGGTAG 480		
QY	1820	ACTTTAAGCCCTGTTTAAAAACATTTCTTACTTAACTTTTGCAAAATTAACCCCATTTG 1878		
Db	481	CCTAACCGGTAAATATACATAGTAG 504		
QY	1879	CTTACCTGTAAATATACATAGTAG 1901		

RESULT	5	AA513750	473 bp	mRNA	EST	20-AUG-1997
LOCUS		nh89g03.s1	NCI_CGAP_Brl.1	Homo sapiens	cdNA clone	IMAGE:965620.
DEFINITION		mRNA sequence.				
ACCESSION		AA513750				
NID		g2252162				
KEYWORDS		EST.				
SOURCE		human.				
ORGANISM		Homo sapiens				
		Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;				
		Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea;				
		Homoe.				
REFERENCE		1 (bases 1 to 473)				
AUTHORS		NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap .				
TITLE		National Cancer Institute, Cancer Genome Anatomy Project (CGAP)				
JOURNAL		Tumor Gene Index				
COMMENT		Unpublished (1997)				

Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
cDNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/dbbr/image/image.html

Insert Length: 2969 Std Error: 0.00
Seq primer: -40m13 fwd. ET from Amersham

FEATURES
source

169	a	57	C	91	g	156	t
							mRNA
							BASE COUNT
							ORIGIN

Query Match	23.1%	Score 468;	DB 8;	Length 473;
Best Local Similarity	99.6%;	Pred. No. 0.00e+00;		
Matches	470;	Conservative 0;	Mismatches 0;	Indels 0; Gaps 0;
Db	1	AGGTAAACAGGATGTAAGTGTATATACACAGAAATAATGTTTATCTCGAAATATTTACA	60	
Cp	2002	AGGTAAACAGGATGTAAGTGTATATACACAGAAATAATGTTTATCTCGAAATATTTACA	1943	
Db	61	GTGTTGGTTAAAGCAATATTTTACAACTTTTAAAGGTAAACTACTATGTATATTACAGG	120	
Cp	1942	GTGTTGGTTAAAGCAATATTTTACAACTTTTAAAGGTAAACTACTATGTATATTACAGG	1883	
Db	121	TAAAGTCACAAATGGTTTAAATTTGCAAAAGCTTAAGTAAGAAATGTTTTAAACAAGGCTTAA	180	
Cp	1882	TAAAGTCACAAATGGTTTAAATTTGCAAAAGCTTAAGTAAGAAATGTTTTAAACAAGGCTTAA	1823	
Db	181	AGTACTCAAGTCAATTTATAAAATTTATATCTTTTGCCTTTAACTTTGAGAAATCATGCTA	240	
Cp	1822	AGTACTCAAGTCAATTTATAAAATTTATATCTTTTGCCTTTAACTTTGAGAAATCATGCTA	1763	
Db	241	TAGAAATGGTTTAATGTGCTTCTAAATAATGGAAGTATTGTAGCTGGAATGTGATACATGT	300	
Cp	1762	TAGAAATGGTTTAATGTGCTTCTAAATAATGGAAGTATTGTAGCTGGAATGTGATACATGT	1703	
Db	301	AACAGTTTTAAGTTCCCATTTCAAGGTATAAAATGATGAATTTGTTGTAAGACTTTAGACACTG	360	
Cp	1702	AACAGTTTTAAGTTCCCATTTCAAGGTATAAAATGATGAATTTGTTGTAAGACTTTAGACACTG	1643	
Db	361	AGTCTCCGTCTGGAGCTGATGAAGATGTTTGAGATAACAGCCAGCTTTTCTCAACAGGGT	420	
Cp	1642	AGTCTCAGTCTGGAGCTGATGAAGATGTTTGAGATAACAGCCAGCTTTTCTCAACAGGGT	1583	
Db	421	TTGTGACCCACAAGTTTGGGCCCCAGAGAAAATTTGAAGCAATTTCCATGTGA	472	
Cp	1582	TTGTGACCCACAAGTTTGGGCCCCAGAGAAAATTTGAAGCAATTTCCATGTGA	1531	

RESULT	6	AA857092	474 bp	mRNA	EST	09-MAR-1998				
LOCUS		af62a09.s1	NCI_CGAP_Co8	Homo sapiens	cDNA clone	IMAGE:1434904 3',				
DEFINITION		mRNA sequence.								
ACCESSION		AA857092								
NID		92945394								
KEYWORDS		EST.								
SOURCE		human.								
ORGANISM		Homo sapiens								
		Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;								
		Primates; Catarrhini; Hominoidea; Homo.								
REFERENCE		1 (bases 1 to 474)								
AUTHORS		NCI-CGAP	http://www.ncbi.nlm.nih.gov/ncicgap.							
TITLE		National Cancer Institute, Cancer Genome Anatomy Project (CGAP),								
		Tumor Gene Index								
JOURNAL		Unpublished (1997)								
COMMENT		Contact: Robert Strausberg, Ph.D.								
		Tel: (301) 496-1550								
		Email: Robert.Strausberg@nih.gov								
		Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.								
		Emmert-Buck, M.D., Ph.D.								
		cDNA Library Preparation: M. Bento Soares, Ph.D.								
		DNA Library Arrayed by: Greg Lennon, Ph.D.								
		Sequencing by: Washington University Genome Sequencing Center								
		Clone distribution: NCI-CGAP clone distribution information can be								
		found through the I.M.A.G.E. Consortium/LLNL at:								
		www-bio.llnl.gov/bbrp/image/image.html								
		Seq primer: -40ml3 fwd. ET from Amersham								
		High quality sequence stop: 464.								
FEATURES		Location/Qualifiers								
source		1..474								
		/organism="Homo sapiens"								
		/note="Organ: colon; Vector: pT7T3D-Pac (Pharmacia) with a								
		modified polylinker; 1st strand cDNA was prepared from								
		colon adenocarcinoma, and was then primed with a Not I -								
		oligo(dT) primer. Double-stranded cDNA was ligated to Eco								
		RI adaptors (Pharmacia), digested with Not I and cloned								
		into the Not I and Eco RI sites of the modified pT7T3								
		vector. Library is normalized. Library was constructed by								
		Bento Soares and M. Fatima Bonaldo."								
		/db_xref="taxon:9606"								
		/clone="IMAGE:1434904"								
		/clone_lib="NCI_CGAP_Co8"								
		/tissue_type="adenocarcinoma"								
		/lab_host="DH10B"								
BASE COUNT		171 a	54 c	92 g	157 t					
ORIGIN										
		Query Match 23.1%; Score 468; DB 15; Length 474;								
		Best Local Similarity 99.8%; Pred.No. 0.00e+00;								
		Matches 474; Conservative 0; Mismatches 0; Indels 1; Gaps 1;								
Db	1	AGGTAAACAGGATGTTAAAGTTTATATACAGAATAATAATCTTCTGAAATATTACA 60								
Cp	2002	AGGTAAACAGGATGTTAAAGTTTATATACAGAATAATAATCTTCTGAAATATTACA 1943								
Db	61	GTGTGGTTAAAGCAATATTTTACAACCTTTTAAAGGTTAACTACTATGTATATTACAGG 120								
Cp	1942	GTGTGGTTAAAGCAATATTTTACAACCTTTTAAAGGTTAACTACTATGTATATTACAGG 1883								
Db	121	TAAGCT								

Accession	Sequence
Cp 1762	TAGAAATGGTTAATGGCTTCTTAATAAATGGAAGTATTGTAGCTGGAAATGTGATACATGT 1700
Db 301	AACAGTTTAAAGTTCCCATTTGAAGGTATATAAATGATGAATTTGTTAAGACTTTAGACACTG 360
Cp 1702	AACAGTTTAAAGTTCCCATTTGAAGGTATATAAATGATGAATTTGTTAAGACTTTAGACACTG 1643
Db 361	AGTCTCAGTCTGGAGCTGATGAAGATGTTTGAGATTAACAGCCAGCTTTATCTCAACAGGGT 420
Cp 1642	AGTCTCAGTCTGGAGCTGATGAAGATGTTTGAGATTAACAGCCAGCTTTATCTCAACAGGGT 1583
Db 421	TTGTGACC-ACAAGTTTGGCCACAGAGAAAATTGAAGCAATTTGTCATGTTATGA 474
Cp 1582	TTGTGACCCACAAGTTTGGCCACAGAGAAAATTGAAGCAATTTGTCATGTTATGA 1528

RESULT	7	W72225	460 bp	mrna	EST	17-OCT-1996
LOCUS	2d7oe11.s1	Soares fetal heart	NBH19W	Homo sapiens	cDNA clone	
DEFINITION	346028	3', mRNA sequence.				
ACCESSION	W72225					
NID	91382674					
KEYWORDS	EST.					
SOURCE	human.					
ORGANISM	Homo sapiens					
Eukaryotes:	mitochondrial eukaryotes; Metazoa: Chordata;					
Vertebrata:	Eutheria: Primates; Catarrhini: Hominoidea: Homo.					
REFERENCE	1 (bases 1 to 460)					
AUTHORS	Hillier, L., Clark, N., Dubucque, T., Elliston, K., Hawkins, M., Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Maier, M., Parsons, J., Rifkin, L., Rohlfing, T., Soares, M., Tan, F., Trevisan, E., Waterston, R., Williamson, A., Wohlmann, P. and Wilson, R.					
TITLE	The WashU-Merck EST Project					
JOURNAL	Unpublished (1995)					
COMMENT	Contact: Wilson RK WashU-Merck EST Project Washington University School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108 Tel: 314 286 1800 Fax: 314 286 1810 Email: est@wustl.wustl.edu This clone is available royalty-free through LLNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information. Insert Length: 1550 Std Error: 0.00 Seq primer: mob.REGA+ET High quality sequence stop: 384.					

FEATURES	source
1..460	Location/Qualifiers
/organism="Homo sapiens"	
/note="Organ: heart; Vector: pT73D (Pharmacia) with a modified polylinker; Site1: Not I; Site 2: Eco RI; ls strand cDNA was primed with a Not I - oligo(dT) primer: TGTTCACATCTGAAGTGGAGCGCCGATCTTTTATTTT 3']	
double-stranded cDNA was size selected, ligated to Eco adapters (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified pT73 vector (Pharmacia). Library went through one round of normalization to a Cot = 5. Library constructed by M.Fatima Bonaldo. This library was constructed from the same fetus as the fetal lung library. Soares fetal lung NBH19W."	
/db_xref="taxon:9606"	
/clone="346028"	
/clone_lib="Soares fetal heart NBH19W"	
/sex="unknown"	
/dev_stage="19 weeks"	
/lab_host="DH10B (ampicillin resistant)"	
complement(<1..>460)	
164 a	55 c 87 g 154 t

BASE COUNT	ORIGIN
Query Match	22.3%; Score 453; DB 23; Length 460;

<p>/organism="Homo sapiens"</p> <p>/note="Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was prepared from mRNA obtained from pooled 8-9 week</p>	<p>Tumor Gene Index</p> <p>Unpublished (1997)</p> <p>JOURNAL</p> <p>COMMENT</p> <p>Contact: Robert Strausberg, Ph.D.</p>
--	--

Tel: (301) 496-1550
 Email: Robert.Strausberg@nih.gov
 Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.
 CDNA Library Preparation: M. Bento Soares, Ph.D.
 CDNA Library Arrayed by: Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CCAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
 www-bio.llnl.gov/bbrp/image/image.html

Insert Length: 870 Std Error: 0.00
 Seq primer: -40ml3 fwd. ET from Amersham
 High quality sequence stop: 377.

FEATURES

Location/Qualifiers
 1..482
 /organism="Homo sapiens"
 /note="Organ: colon; Vector: p773D-Pac (Pharmacia) with a modified polylinker; 1st strand cDNA was prepared from colon adenocarcinoma, and was then primed with a Not I - oligo(dT) primer. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified p773 vector. Library is normalized. Library was constructed by Bento Soares and M. Fatima Bonaldo."
 /db_xref="taxon:9606"
 /clone="IMAGE:1457445"
 /clone_lib="NCI-CCAP_Co8"
 /tissue_type="adenocarcinoma"
 /lab_host="DH10B"

BASE COUNT 103 a 119 c 127 g 133 t

Query Match 21.4%; Score 433; DB 18; Length 482;
 Best Local Similarity 99.8%; Pred. No. 0.00e+00;
 Matches 434; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 10 AAAATATTCATGATTTTATAGTTTGAATATTTCTACAAGATTCGGGTGGGCTTTTCCTT 69
 Cp 1406 AAAATATTCATGATTTTATAGTTTGAATATTTCTACAAGATTCGGGTGGGCTTTTCCTT 1347

Db 70 TAGGTGAAACAGCTATCCACTCCTGTGGCCTTATAACTCAGAAATGCTGGGATGCAA 129
 Cp 1346 TAGGTGAAACAGCTATCCACTCCTGTGGCCTTATAACTCAGAAATGCTGGGATGCAA 1287

Db 130 AGTGCAGGAGCGGGGAGCTGCCAGCTGAGACTGGAGCAGCTAGGAGTGTGCTT 189
 Cp 1286 AGTGCAGGAGCGGGGAGCTGCCAGCTGAGACTGGAGCAGCTAGGAGTGTGCTT 1227

Db 190 GGGAGACGGAGCTGAGATCCGGAGCAGAAATGGTCAGCCGTCTCTGGAGCAGGCCCTG 249
 Cp 1226 GGGAGACGGAGCTGAGATCCGGAGCAGAAATGGTCAGCCGTCTCTGGAGCAGGCCCTG 1167

Db 250 TCGGAGCCATCAGGATGCGGGAGCTAGCACTGCAGCTTCGGAATGCTCGGGAGATGCGC 309
 Cp 1166 TCGGAGCCATCAGGATGCGGGAGCTAGCACTGCAGCTTCGGAATGCTCGGGAGATGCGC 1107

Db 310 TTGAACCTCTCTGCCCTTCTGCCACCCCTTCACCGAGGTGATCACCAGCTCCCCACCC 369
 Cp 1106 TTGAACCTCTCTGCCCTTCTGCCACCCCTTCACCGAGGTGATCACCAGCTCCCCACCC 1047

Db 370 TGTTCCTGTCACCATGACAGATAGGCGCGTTGATGCTGCTTCATCTCCTCAGGTGCA 429
 Cp 1046 TGTTCCTGTCACCATGACAGATAGGCGCGTTGATGCTGCTTCATCTCCTCAGGTGCA 987

Db 430 TCGAAGCTGCTTTG 444
 Cp 986 TCGAAGCTGCTTTG 972

RESULT 10
 LOCUS AA431859 427 bp mRNA EST 22-MAY-1997
 DEFINITION zw51g08.s1 Soares total fetus Nb2HF8 9w Homo sapiens CDNA clone

773630 3', mRNA sequence.
 AA431859
 NID G2115567
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 427)
 AUTHORS Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S., Kucaba,T., Lacy,M., Le.N., Lennon,G., Marra,M., Martin,J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F., Theising,B., White,Y., Wylie,T., Waterston,R. and Wilson,R.
 WashU-Merck EST Project 1997
 Unpublished (1997)

CONTACT: Wilson RK
 WashU-Merck EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@watson.wustl.edu
 This clone is available royalty-free through LLNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
 Seq primer: -40ml3 fwd. ET from Amersham
 High quality sequence stop: 214.

Location/Qualifiers
 1..427

/organism="Homo sapiens"
 /note="Vector: p773D-Pac (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA was prepared from mRNA obtained from pooled 8-9 week (total) fetus material with a Not I oligo(dT) primer [5' TGTTACCAATCTGAAGTGGGAGCGCGCTTAATTTTTTTTTTTT 3']
 Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified p773 vector. Library went through one round of normalization, and was constructed by Bento Soares and M. Fatima Bonaldo."
 /db_xref="taxon:9606"
 /clone="773630"
 /dev_stage="8-9 weeks"
 /lab_host="DH10B"
 complement(<1..>427)
 154 a 46 c 79 g 148 t

FEATURES

Location/Qualifiers
 1..427
 /organism="Homo sapiens"
 /note="Vector: p773D-Pac (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA was prepared from mRNA obtained from pooled 8-9 week (total) fetus material with a Not I oligo(dT) primer [5' TGTTACCAATCTGAAGTGGGAGCGCGCTTAATTTTTTTTTTTT 3']
 Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified p773 vector. Library went through one round of normalization, and was constructed by Bento Soares and M. Fatima Bonaldo."
 /db_xref="taxon:9606"
 /clone="773630"
 /dev_stage="8-9 weeks"
 /lab_host="DH10B"
 complement(<1..>427)
 154 a 46 c 79 g 148 t

Query Match 21.0%; Score 425; DB 5; Length 427;
 Best Local Similarity 99.8%; Pred. No. 0.00e+00;
 Matches 426; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 1 TTTTATAGTAAACAGGATGTAAGTTTATATACAGAATAATATGTTTCTCAATA 60
 Cp 2008 TTTTATAGTAAACAGGATGTAAGTTTATATACAGAATAATATGTTTCTCAATA 1949

Db 61 TTTACAGTGTGGTTAAAGCAATATTTTACAACCTTTTAAAGGTAACACTATATAT 120
 Cp 1948 TTTACAGTGTGGTTAAAGCAATATTTTACAACCTTTTAAAGGTAACACTATATAT 1889

Db 121 TACAGGTAAAGTACAAATGGGTTTAAATTTGCAAAAGTTAAAGTAAGAAATGTTTAAACAAG 180
 Cp 1888 TACAGGTAAAGTACAAATGGGTTTAAATTTGCAAAAGTTAAAGTAAGAAATGTTTAAACAAG 1829

Db 181 GCTTAAAGTACTCAAGTCAATTAATAATTTATATATCTTTTGCCTTTTACTTGAAGAAATC 240
 Cp 1828 GCTTAAAGTACTCAAGTCAATTAATAATTTATATATCTTTTGCCTTTTACTTGAAGAAATC 1769

Db 241 ATGCTATAGAAATGGTTAATGTGCTTCTTAATAAATGGAAGTATTTAGCTGGAATGTGAT 300
 Cp 1768 ATGCTATAGAAATGGTTAATGTGCTTCTTAATAAATGGAAGTATTTAGCTGGAATGTGAT 1709

Db 301 ACATGTAACAGTTTAAAGTTCCATTGAAGGTATATAAAATGATGAATTTGTTAAGACTTAG 360
 |||||
 Cp 1708 ACATGTAACAGTTTAAAGTTCCATTGAAGGTATATAAAATGATGAATTTGTTAAGACTTAG 1649
 |||||
 Db 361 ACACGTAGTCTCAGTCTGGAGCTGATGAAGATGTTGAGATACAGCCAGCTTTATCTCAA 420
 |||||
 Cp 1648 ACACGTAGTCTCAGTCTGGAGCTGATGAAGATGTTGAGATACAGCCAGCTTTATCTCAA 1589
 |||||
 Db 421 CAGGGTT 427
 |||||
 Cp 1588 CAGGGTT 1582
 |||||

RESULT 11 AA927991 518 bp mRNA EST 22-APR-1998
 LOCUS on47d08.s1 NCI_CGAP_Co8 Homo sapiens cDNA clone IMAGE:1559823 3'
 DEFINITION similar to TR:008862 008862 SECRETED FRIZZLED RELATED PROTEIN
 SFRP-2. [2] TR:P97299 ; mRNA sequence.

ACCESSION AA927991
 NID g3076735
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
 Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 518)
 AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 JOURNAL Unpublished (1997)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Tel: (301) 496-1550
 Email: Robert.Strausberg@nih.gov
 Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
 Emmert-Buck, M.D., Ph.D.
 cDNA Library Preparation: M. Bento Soares, Ph.D.
 DNA Library Arrayed by: Greg Lennon, Ph.D.
 Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 www-bio.llnl.gov/bbrp/image/image.html

Seg primer: -40m13 fwd. ET from Amersham
 High quality sequence stop: 194.
 FEATURES
 source Location/Qualifiers
 1..518
 /organism="Homo sapiens"
 /note="Organ: colon; Vector: pT7T3D-Pac (Pharmacia) with a
 modified polylinker; 1st strand cDNA was prepared from
 colon adenocarcinoma, and was then primed with a Not I -
 oligo(dT) primer. Double-stranded cDNA was ligated to Eco
 RI adaptors (Pharmacia), digested with Not I and cloned
 into the Not I and Eco RI sites of the modified pT7T3
 vector. Library is normalized. Library was constructed by
 Bento Soares and M. Fatima Bonaldo."
 /db_xref="taxon:9606"
 /clone="IMAGE:1559823"
 /clone_lib="NCI_CGAP_Co8"
 /tissue_type="adenocarcinoma"
 /lab_host="DH10B"
 BASE COUNT 112 a 123 c 141 g 142 t
 ORIGIN

Query Match 20.7%; Score 420; DB 17; Length 518;
 Best Local Similarity 98.8%; Pred.No. 0.00e+00;
 Matches 425; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Db 10 AAAATATTTCATGATTTTATTAGTTTGAATATTTCTACAAGATTCGGGTGGGCTTTTCCTT 69
 |||||
 Cp 1406 AAAATATTTCATGATTTTATTAGTTTGAATATTTCTACAAGATTCGGGTGGGCTTTTCCTT 1347
 |||||

Db 70 TAGTGAAGAACAGCTATCCACTCTCTGTGGCTTTATAACTCAGGAATGCTGGGGATGCAA 129
 |||||
 Cp 1346 TAGTGAAGAACAGCTATCCACTCTCTGTGGCTTTATAACTCAGGAATGCTGGGGATGCAA 1287
 |||||
 Db 130 AGTGTGAAAAGGAGGGGGAAGCTGCCAGGCTGAGAGCTGAGAGCTAGGAGTGTGCTT 189
 |||||
 Cp 1286 ACCTGTGAAAAGGAGGGGGAAGCTGCCAGGCTGAGAGCTGAGAGCTAGGAGTGTGCTT 1227
 |||||
 Db 190 GGGAGCGGAGCTGAGATCCCGAGCAGAGAAATGGTCAGCCGTGCTGTGAGCAGGCCTG 249
 |||||
 Cp 1226 GGGAGCGGAGCTGAGATCCCGAGCAGAGAAATGGTCAGCCGTGCTGTGAGCAGGCCTG 1167
 |||||
 Db 250 TCGAGCCATCAGGATGCCGGGACTAGCACTGCAGCTTGGGATGCTGCGGGAGATGCCG 309
 |||||
 Cp 1166 TCGAGCCATCAGGATGCCGGGACTAGCACTGCAGCTTGGGATGCTGCGGGAGATGCCG 1107
 |||||
 Db 310 TTGAATCTCTGCCCCCTTCTGCCACCGCTTCGACGAGGTATCACCAGCTCCCAACC 369
 |||||
 Cp 1106 TTGAATCTCTGCCCCCTTCTGCCACCGCTTCGACGAGGTATCACCAGCTCCCAACC 1047
 |||||
 Db 370 TGTTCCTGTCCTCCATGACAGATAGGCGCGCTTATGTCTTCATCTACTACAGGTGCAC 429
 |||||
 Cp 1046 TGTTCCTGTCCTCCATGACAGATAGGCGCGCTTATGTCTTCATCTACTACAGGTGCAC 987
 |||||

RESULT 12
 LOCUS N64004 446 bp mRNA EST
 DEFINITION YY64d09.s1 Homo sapiens cDNA clone 278321 3'.
 ACCESSION N64004
 NID g1211833
 KEYWORDS EST.
 SOURCE human clone=278321 primer=ml3 -40 forward library=Soares multiple
 sclerosis 2NBHSP vector=pT7T3D (Pharmacia) with a modified
 polylinker V-TYPE: phagemid host=DH10B (ampicillin resistant)
 Rsitel=Not I Rsitel2=Eco RI 46 year old male. 1st strand cDNA was
 primed with a Not I - oligo(dT) primer
 [5'-TGTTACCAATCTGAAGTGGGAGGCGCGCATTTTCTTTTCTTTT-3'],
 double-stranded cDNA was size selected, ligated to Eco RI adaptors
 (Pharmacia), digested with Not I and cloned into the Not I and Eco
 RI sites of a modified pT7T vector (Pharmacia). Library went
 through one round of normalization to a Cot = 5. Library
 constructed by Bento Soares and M. Fatima Bonaldo. RNA from 4
 multiple sclerosis lesions from one patient was kindly provided by
 Dr. Kevin G. Becker (NINDS/NIH).

ORGANISM Homo sapiens
 Eukaryota; Metazoa; Eumetazoa; Bilateria; Coelomata;
 Deuterostomia; Chordata; Vertebrata; Gnathostomata; Osteichthyes;
 Sarcopterygii; Chonata; Tetrapoda; Amniota; Mammalia; Theria;
 Eutheria; Archonta; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 446)
 Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M.,
 Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M.,
 Parsons, J., Rifkin, L., Rohlfing, T., Soares, M., Tan, F.,
 Trevas, E., Waterston, R., Williamson, A., Wohldmann, P. and
 Wilson, R.
 The WashU-Merck EST Project
 Unpublished (1995)

REFERENCE
 AUTHORS
 The WashU-Merck EST Project
 Unpublished (1995)

TITLE
 JOURNAL
 COMMENT

Contact: Wilson RK
 WashU-Merck EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@wustl.edu
 High quality sequence stops: 225
 Source: IMAGE Consortium, LLNL
 This clone is available royalty-free through LLNL; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.

FEATURES	source	Location/Qualifiers
		1. .446
		/organism="Homo sapiens"
		/clone="278321"
	mRNA	<1. .>446
BASE COUNT		157 a 51 c 87 g 14
ORIGIN		

Query Match 20.5%; Score 416; DB 21; Length 446;
Best Local Similarity 98.9%; pred. No. 0.00e+00;
Matches 440; Conservative 0; Mismatches 1; Indels

Db	2	TAGCTAAACAGGATGTAAAGTTTATATACAGAATAATAATGTTTATCTGAAATATTTAC	61
Cp	2003	TAGCTAAACAGGATGTAAAGTTTATATACAGAATAATAATGTTTATCTGAAATATTTAC	1944
Db	62	AGTGTGTGGTTAAAGCAATATTTTACAACTTTTAAAGGTAAACTACTATGTATATTACAG	121
Cp	1943	AGTGTGTGGTTAAAGCAATATTTTACAACTTTTAAAGGTAAACTACTATGTATATTACAG	1884
Db	122	GTAAGCTACAATGGGTTTAAITTCGAAAAGTTAAGTAAGAAATGTTTTAAACAAGGCTTA	181
Cp	1883	GTAAGCTACAATGGGTTTAAITTCGAAAAGTTAAGTAAGAAATGTTTTAAACAAGGCTTA	1824
Db	182	AGTACTCAAGTCAAAATATATAAATTTATATCTTTTGCCTTTTACTTGAAGAAATCATGCT	241
Cp	1823	AGTACTCAAGTCAAAATATATAAATTTATATCTTTTGCCTTTTACTTGAAGAAATCATGCT	1764
Db	242	ATAGAAATGGTTAATGTGCTTCCTAAATAAATGGAAGTATTGTAGCTGGAAATGTGATACATG	301
Cp	1763	ATAGAAATGGTTAATGTGCTTCCTAAATAAATGGAAGTATTGTAGCTGGAAATGTGATACATG	1704
Db	302	TACAGCTTTAAGTTCCTCCATTGAAGGTATAAAATNGATGAATTTGTTGAAGACCTTTAGACA	361
Cp	1703	TACAGCTTTAAGTTCCTCCATTGAAGGTATAAAAT-GATGAATTTGTTGAAGAC-TTAGACA	1646
Db	362	CTGGAGTCTCAGTCTGGGAGCTCATGAAGATGTTGAGATAACAGCCAGCTTTATCTCAAC	421
Cp	1645	CTG-AGTCTCAGTCTGG-AGCTGATGNAGATGTTGAGATAACAGCCAGCTTTATCTCAAC	1588
Db	422	AGGTTTGTGACCCNCNAAGTTTGGG	446
Cp	1587	AGGTTTGTGACCCACAAGTTTGGG	1563

[illegible]

Contact: Wilson RK
WashU-Merck EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LLNL ; contact the

Email: estewartson.wustl.edu
This clone is available royalty-free through LLNL ; contact the

IMAGE Consortium (infoimage.llnl.gov) for further information.
Insert Length: 1550 Std Error: 0.00
Seq primer: mob.REGA+ET
High quality sequence stop: 424.

FEATURES
source

```

1: 1042
/organism="Homo sapiens"
/note="Organ: heart; Vec:
modified polylinker; Sit:
strand cDNA was primed w
TGTTACCAATCTCAAGTGGGAGCG
double-stranded cDNA was
adapters (Pharmacia), di
the Not I and Eco RI sit
(Pharmacia). Library wen
normalization to a Cot =
M.Fatima Bonaldo. This l
same fetus as the fetal
NBHL19W."
/db_xref="taxon:9606"
/clone="346028"
/clone_lib="Soares fetal
/sex="unknown"
/dev_stage="19 weeks"
/lab_host="DH10B (ampicid
<1. >462
128 a 133 c 106 g 9
mRNA
BASE COUNT
ORIGIN

```

	Query Match	20.4%	Score 413:	DB 23:	Length 462:
	Best Local Similarity 98.0%:		Pred. No. 0.00e+00:		
	Matches 449:	Conservative 0:	Mismatches 3:	Indels 6:	Gaps 6:
Db	6	TCGCTCTTCGCCCGCGCTCTGCCCTCGATGACCTAGACGAGACCATCCAGCATGCCACTCG	65		
Qy	544	TCGCTCTTCGCCCGCGCTCTGCCCTCGATGACCTAGACGAGACCATCCAGCATGCCACTCG	603		
Db	66	CTCTGN-TGCAGGTGAAGGACCGCTGCGCCCGGTCATGTCCGCCCTTCGN-TTCCCGCTGG	123		
Qy	604	CTTCGCTGCAGGTGAAGGACCGCTGCGCCCGGTCATGTCCGCCCTTCGSGCTTCCCGCTGG	663		

DB	124	CCGAGATGCTT	GAGTGCAGCGTTTCC	CCCAAGCAACGACCTTTG	CAATCCCCCTCCGC	181
DB	125	CCGAGATGCTT	GAGTGCAGCGTTTCC	CCCAAGCAACGACCTTTG	CAATCCCCCTCCGC	182
DB	126	CCGAGATGCTT	GAGTGCAGCGTTTCC	CCCAAGCAACGACCTTTG	CAATCCCCCTCCGC	183
Qy	664	CCGAGATGCTT	GAGTGCAGCGTTTCC	CCCAAGCAACGACCTTTG	CATCCCCCTC	722
DB	184	TACGAGN	-ACCACCTCCTGC	CAGCACACCGAGGAAGCTCC	CAAGGCTGTG	242
DB	185	TACGAGN	-ACCACCTCCTGC	CAGCACACCGAGGAAGCTCC	CAAGGCTGTG	243
Qy	723	TAGCAGG	CCACCTCCTGC	CAGCACACCGAGGAAGCTCC	CAAGGCTGTG	782
DB	243	AAATAAAATGATGATG	ACAACGACATAATG	GAAACGCTTTGT	AAAAATGATTTTGC	302
DB	244	AAATAAAATGATGATG	ACAACGACATAATG	GAAACGCTTTGT	AAAAATGATTTTGC	303
Qy	783	AAATAAAATGATGATG	ACAACGACATAATG	GAAACGCTTTGT	AAAAATGATTTTGC	842
DB	303	GAATAAAAGTGAAGGAGAT	AACCTACATCAAC	CGAGATACCAAAATCATCTCTGG	AAGA	362
Qy	843	GAATAAAAGTGAAGGAGAT	AACCTACATCAAC	CGAGATACCAAAATCATCTCTGG	AAGA	901
DB	363	CCAAGAGCAAGACCATTT	ACAAGCTGAACGGTGTG	TCCGAAAGGAGCACTTGA	AGAAATCG	422
Qy	902	CCAAGAGCAAGACCATTT	ACAAGCTGAACGGTGTG	TCCGAAAGGAGCACTTGA	AGAAATCG	960
DB	423	GTGCTGTGGCTCA	AGACAGCTTGC	AGTGCACCTGTGA	460	
Qy	961	GTGCTGTGGCTCA	AGACAGCTTGC	AGTGCACCTGTGA	998	

RESULT	14
LOCUS	AA449749
DEFINITION	z07e10.s1 Soares total fetus Nb2H8 9w Homo sapiens cDNA clone
ACCESSION	785802 3', mRNA sequence.
NID	AA449749
KEYWORDS	EST.
	EST
	409 bp
	mRNA
	04-JUN-1997

SOURCE human.
ORGANISM Homo sapiens
Eukaryotes; mitochondrial eukaryotes; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.

REFERENCE 1 (bases 1 to 409)
AUTHORS Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisel, G., Jost, S., Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M., Martin, J., Moore, B., Schellenberg, K., Steptoe, M., Tan, F., Theising, B., White, Y., Wylie, T., Waterston, R., and Wilson, R.
TITLE WashU-Merck EST Project 1997
JOURNAL Unpublished (1997)
COMMENT Contact: Wilton RK
WashU-Merck EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LLNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -4Jm13 fwd. Et from Amersham
High quality sequence stop: 408.

FEATURES
source
1..409
/organism="Homo sapiens"
/note="vector: p7T3D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was prepared from mRNA obtained from pooled 8-9 week (total) fetus material with a Not I - oligo(dT) primer [5' TGTACCAATCTGAAGTGGAGCGCGCTTAATTTTTTTTTTTT 3']
TGTACCAATCTGAAGTGGAGCGCGCTTAATTTTTTTTTTTT 3']
Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified p7T3 vector. Library went through one round of normalization, and was constructed by Bento Soares and M. Fatima Bonaldo."
/db_xref="taxon:9606"
/clone_lib="Soares total fetus NB2HF8 9w"
/dev_stage="8-9 weeks"
/lab_host="DH10B"
complement(<1..>409)
/db_xref="GDB:5963421" 144 t

mRNA
BASE COUNT 148 a 42 c 75 g 144 t
ORIGIN

Query Match 20.2%; Score 409; DB 7; Length 409;
Best Local Similarity 100.0%; Pred. No. 0.00e+00;
Matches 409; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 TTTTCTTTAGTAAACAGGATGTAAGTTTATATACAGAATATAATGTTTATCTGAAA 60
Cp 2010 TTTTCTTTAGTAAACAGGATGTAAGTTTATATACAGAATATAATGTTTATCTGAAA 1951
Db 61 TATTACAGTGTGGTTAAAGCATATTTTACAACTTTTAAAGGTAACACTACTATGAT 120
Cp 1950 TATTACAGTGTGGTTAAAGCATATTTTACAACTTTTAAAGGTAACACTACTATGAT 1891
Db 121 ATTACAGTAACTACAACTGGTTTAAATTTGCAAAAGTTAAGTAAGAAATGTTTAAACA 180
Cp 1890 ATTACAGTAACTACAACTGGTTTAAATTTGCAAAAGTTAAGTAAGAAATGTTTAAACA 1831
Db 181 AGGCTTAAAGTAACTAACTCAATTTATATATATATATCTTTTGCCTTTTACTTGAAGAAA 240
Cp 1830 AGGCTTAAAGTAACTAACTCAATTTATATATATATATCTTTTGCCTTTTACTTGAAGAAA 1771
Db 241 TCATGCTATAGAAATGTTTAAATGCTTCTTAATAAATGGAAGTATGCTAGCTGAATGTC 300
Cp 1770 TCATGCTATAGAAATGTTTAAATGCTTCTTAATAAATGGAAGTATGCTAGCTGAATGTC 1711
Db 301 ATACATGTAACAGTTTAAAGTTCCTCCATTTGAAAGGTATATAAATGATGAATGTTGTAAGACTT 360

Cp 1710 ATACATGTAACAGTTTAAAGTTCCTCCATTTGAAGGTATATAAATGATGAATGTTGTAAGACTT 1651
Db 361 AGACACTGAGTCTCAGTCTCGAGCTGATGAAGATGTTGAGATAACAGCC 409
Cp 1650 AGACACTGAGTCTCAGTCTCGAGCTGATGAAGATGTTGAGATAACAGCC 1602

RESULT 15
LOCUS N32415 445 bp mRNA EST 10-JAN-1996
DEFINITION YW83C06.s1 Homo sapiens cDNA clone 258826 3'
ACCESSION N32415
NID g1152814
KEYWORDS EST.
SOURCE human clone=258826 primer=ml3 -40 forward library=Soares placenta 8to9weeks 2NbHP8to9W vector=p7T3D (Pharmacia) with a modified polylinker host=DH10B (ampicillin resistant) Rsitel-Not I Rsitel2=Eco RI two placentae: one from 8 weeks and another from 9 weeks post conception. 1st strand cDNA was primed with a Not I - oligo(dT) primer
[5'-TGTACCAATCTGAAGTGGAGCGCGCGATTTTTTTTTTTTTTTT-3'], double-stranded cDNA was size selected, ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified p7T3 vector (Pharmacia). Library constructed by Bento Soares and M. Fatima Bonaldo.

ORGANISM Homo sapiens
Eukaryotes; Metazoa; Eumetazoa; Bilateria; Coelomata; Deuterostomia; Chordata; Vertebrata; Gnathostomata; Osteichthyes; Sarcopterygii; Choonata; Tetrapoda; Amniota; Mammalia; Theria; Eutheria; Archonta; Primates; Catarrhini; Hominoidea; Homo. 1 (bases 1 to 445)
Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M., Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M., Parsons, J., Rifkin, L., Rohlfing, T., Soares, M., Tan, F., Trevaskis, E., Waterston, R., Williamson, A., Wohlmann, P., and Wilson, R.
TITLE The WashU-Merck EST Project
JOURNAL Unpublished (1995)
COMMENT Contact: Wilton RK
WashU-Merck EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
High quality sequence stops: 314
Source: IMAGE Consortium, LLNL
This clone is available royalty-free through LLNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.

FEATURES
source
1..445
/organism="Homo sapiens"
/clone="258826"
BASE COUNT 156 a 55 c 87 g 147 t
ORIGIN

Query Match 20.1%; Score 408; DB 20; Length 445;
Best Local Similarity 98.6%; Pred. No. 0.00e+00;
Matches 434; Conservative 0; Mismatches 2; Indels 4; Gaps 4;

Db 1 AAACAGGATGTAAGTTTATATACAGAATATAATGTTTATCTGAAATATTACAGTGT 60
Cp 1998 AAACAGGATGTAAGTTTATATACAGAATATAATGTTTATCTGAAATATTACAGTGT 1939
Db 61 TGGTTAAACCAATATTTTACAACTTTTAAAGTAACTACTATGTTTATACAGGTAG 120
Cp 1938 TGGTTAAACCAATATTTTACAACTTTTAAAGTAACTACTATGTTTATACAGGTAG 1879
Db 121 CTACATGGGTTTAAATTTGCAAAAGTTAAGTAAGAAATGTTTAAACAGGCTTAAAGTA 180
Cp 1878 CTACATGGGTTTAAATTTGCAAAAGTTAAGTAAGAAATGTTTAAACAGGCTTAAAGTA 1819

```

Db 181 CTCAGTCAATTATAAAATTATATCTTTTGCCTTTTACTTGAAGAAATCATGCTATAGA 240
Cp 1818 CTCAGTCAATTATAAAATTATATCTTTTGCCTTTTACTTGAAGAAATCATGCTATAGA 1759
Db 241 AATGGTTAATGCTCTCTTAATAATGAAGTATTGTAGCTGGGAATGTGATACATGTAACA 300
Cp 1758 AATGGTTAATGCTCTCTTAATAATGAAGTATTGTAGCTGGGAATGTGATACATGTAACA 1699
Db 301 GTTTAAGTTCCCATTTGAAGGTATAAAATGATGAATTGTTGAAGACTTTAGACACTGGAGT 360
Cp 1698 GTTTAAGTTCCCATTTGAAGGTATAAAATGATGAATTGTTGAAGACTTTAGACACTG-AGT 1640
Db 361 CTCAGTCTGGGAGCTGATGAAGATGTTGAGATAACAGCCAGCTTTATCCCCAACAGGGGT 420
Cp 1639 CTCAGTCTGG-AGCTGATGAAGATGTTGAGATAACAGCCAGCTTTATCTC-AACAGGG-T 1583
Db 421 TTGTGACCCACAAGTTTGG 440
Cp 1582 TTGTGACCCACAAGTTTGG 1563

```

Search completed: Wed Aug 5 07:34:29 1998
 Job time : 4465 secs.

THIS PAGE BLANK (section)

Result No.	Query			ID	Description	Pred. No.
	Score	Match	Length			
1	396	17.7	572	W3127	Mouse frizzled-7 prot	2.72e-28
2	394	17.6	685	W31274	Mouse frizzled-8 prot	4.35e-28
3	374	16.7	585	W31271	Human frizzled-5 prot	4.69e-26
4	369	16.5	537	W31270	Mouse frizzled-4 prot	1.51e-25
5	357	16.0	525	W31269	Caenorhabditis frizl	2.45e-24
6	355	15.9	694	W31267	Drosophila frizzled-2	3.92e-24
7	317	14.2	666	W31268	Mouse frizzled-3 prot	2.56e-20
8	284	12.7	709	W31272	Mouse frizzled-6 prot	4.86e-17
9	148	6.6	581	R74187	Chicken p75	3.74e-04
10	124	5.6	605	R74186	Chicken p78	4.55e-02
11	121	5.4	604	W35946	Human netrin-1.	8.16e-02
12	105	4.7	529	R74188	Mouse p78.	1.70e+00
13	100	4.5	732	W26642	Human RECK cancer-inh	4.26e+00
14	99	4.4	804	P253368	H. pylori ORF 09ap114	5.11e+00
15	96	4.3	868	W26610	Rat muscle-specific k	8.79e+00
16	96	4.3	868	W26507	Rat Dmk receptor.	8.79e+00
17	95	4.3	869	W26631	Human musc-specific	1.05e+01
18	95	4.3	869	W26506	Human Dmk receptor.	1.05e+01

Db 51 p is p l o t d i a y n g t i l p n l l i g h t n g e d a g l e v h q f y p l v k v d s p e i r f f l d s m y a p v c 110

QY 44 PANLQCHGIEYQNMRLPNLLGHETMKEVLEQAGAWIPLVMKQCHPDTRKFLCSLFAFVC 103
Db 111 -tylqaiopslkqararqgcaalmnkfgfowperlrcenfovhgageiaygqntsd 167
QY 104 LDDLDETIQCHSLCVQVDRKCAPVNSAFGFPWMDLECDREFP-ODN-DLCIPLASSD 159

RESULT 2
ID W31274 standard; Protein; 585 AA.
AC W31274;
DT 27-APR-1998 (first entry)
DE Mouse frizzled-8 protein Mfz8 (Wnt receptor).
KW Wnt receptor; mouse frizzled-8 protein; Mfz8 gene;
KW signal transduction; cancer; cell growth; cell proliferation.
OS Mus musculus.
PN W09739357-A1.
PD 23-OCT-1997.
PF 11-APR-1997; U06049.
PR 12-APR-1996; US-015307.
PA (UYJO) UNIV JOHNS HOPKINS.
PI Andrew D, Bhanot P, Brink M, Hsieh J, Nathans J,
PI Nusse R, Samos CH, Wangy;
DR N-PSDB; T89889.
PT Identification of Wnt receptor binding modulators - useful for
PT treatment of cancer and growth, development or proliferation related
PT disorders
PS Disclosure; Page 48-50; 61pp; English.
CC This protein comprises the mouse transmembrane receptor,
CC frizzled-8 (Mfz8), encoded by the Mfz8 gene (see T89892). It is
CC an example of a Wnt receptor. Other novel frizzled family members
CC have been identified in human, mouse and Caenorhabditis elegans
CC (see W31268-74) and are considered also to be Wnt receptors. Wnt
CC receptors can be used in a novel, claimed method of screening for
CC compounds which modulate the binding of a Wnt polypeptide (secreted
CC proteins involved in cell-to-cell signalling) to a Wnt receptor.
CC Wnt is involved in (mammary) cancer and other processes involving
CC growth, development and proliferation (both normal and abnormal).
CC Modulators identified by the claimed method are useful for
CC treatment of diseases related to these conditions.
SQ Sequence 685 AA;

Query Match 17.6%; Score 394; DB 27; Length 685;
Best Local Similarity 43.8%; Pred. No. 4.35e-28;
Matches 56; Conservative 23; Mismatches 43; Indels 6; Gaps 6;

Db 40 vplckgiqnytmqngfnhdtqdeaglevhgf-piveiqcsdtkfflcsmytpicle 98
QY 47 LQCHGIEYQNMRLPNLLGHETMKEV-LEQAGAWIPLVMKQCHPDTRKFLCSLFAFVCLD 105

Db 99 dykxlpccrsvcerakagcaplmrtygfawpdrmcdrilpeqgnpdtlcmdyrt-d-lt 157
QY 106 DLDETIQCHSLCVQVDRKCAPVNSAFGFPWMDLECDREFP-ODN-D-LCIPLASSDHL 162

Db 158 taapsppr 165
QY 163 PATEEAPK 170

RESULT 3
ID W31271 standard; Protein; 585 AA.
AC W31271;
DT 27-APR-1998 (first entry)
DE Human frizzled-5 protein Mfz5 (Wnt receptor).
KW Wnt receptor; human frizzled-5 protein; Mfz5 gene;
KW signal transduction; cancer; cell growth; cell proliferation.
OS Homo sapiens.
PN W09739357-A1.
PD 23-OCT-1997.
PF 11-APR-1997; U06049.
PR 12-APR-1996; US-015307.
PA (UYJO) UNIV JOHNS HOPKINS.

PA (STRD) UNIV LELAND STANFORD JUNIOR.
PI Andrew D, Bhanot P, Brink M, Hsieh J, Nathans J,
PI Nusse R, Samos CH, Wangy;
DR N-PSDB; T89889.
PT Identification of Wnt receptor binding modulators - useful for
PT treatment of cancer and growth, development or proliferation related
PT disorders
PS Disclosure; Page 37-39; 61pp; English.
CC This protein comprises the human transmembrane receptor,
CC frizzled-5 (Hfz5), encoded by the Hfz5 gene (see T89889). It is
CC an example of a Wnt receptor. Other novel frizzled family members
CC have been identified in human, mouse and Caenorhabditis elegans
CC (see W31268-74) and are considered also to be Wnt receptors. Wnt
CC receptors can be used in a novel, claimed method of screening for
CC compounds which modulate the binding of a Wnt polypeptide (secreted
CC proteins involved in cell-to-cell signalling) to a Wnt receptor.
CC Wnt is involved in (mammary) cancer and other processes involving
CC growth, development and proliferation (both normal and abnormal).
CC Modulators identified by the claimed method are useful for
CC treatment of diseases related to these conditions.
SQ Sequence 585 AA;

Query Match 16.7%; Score 374; DB 27; Length 585;
Best Local Similarity 38.8%; Pred. No. 4.69e-26;
Matches 47; Conservative 29; Mismatches 41; Indels 4; Gaps 4;

Db 20 lvgraaasakpvcqei-t-vpmcrgigvnlthmpnqfnhdtqdeaglevhgf-wplvei 76
QY 27 LFQOPDFSYKRSNCKPIPANLQCHGIEYQNMRLPNLLGHETMKEV-LEQAGAWIPLVMK 85

Db 77 qcspdlrffictmytpicldpdyhkpccrsvcerakagcplmrygfawpdrmcdril 136
QY 86 QCHPDTRKFLCSLFAFVCLDDETIQCHSLCVQVDRKCAPVNSAFGFPWMDLECDREF 145

Db 137 p 137
QY 146 P 146

RESULT 4
ID W31270 standard; Protein; 537 AA.
AC W31270;
DT 27-APR-1998 (first entry)
DE Mouse frizzled-4 protein Mfz4 (Wnt receptor).
KW Wnt receptor; mouse frizzled-4 protein; Mfz4 gene;
KW signal transduction; cancer; cell growth; cell proliferation.
OS Mus musculus.
PN W09739357-A1.
PD 23-OCT-1997.
PF 11-APR-1997; U06049.
PR 12-APR-1996; US-015307.
PA (UYJO) UNIV JOHNS HOPKINS.
PI Andrew D, Bhanot P, Brink M, Hsieh J, Nathans J,
PI Nusse R, Samos CH, Wangy;
DR N-PSDB; T89889.
PT Identification of Wnt receptor binding modulators - useful for
PT treatment of cancer and growth, development or proliferation related
PT disorders
PS Disclosure; Page 34-35; 61pp; English.
CC This protein comprises the mouse transmembrane receptor,
CC frizzled-4 (Mfz4), encoded by the Mfz4 gene (see T89888). It is
CC an example of a Wnt receptor. Other novel frizzled family members
CC have been identified in human, mouse and Caenorhabditis elegans
CC (see W31268-74) and are considered also to be Wnt receptors. Wnt
CC receptors can be used in a novel, claimed method of screening for
CC compounds which modulate the binding of a Wnt polypeptide (secreted
CC proteins involved in cell-to-cell signalling) to a Wnt receptor.
CC Wnt is involved in (mammary) cancer and other processes involving
CC growth, development and proliferation (both normal and abnormal).
CC Modulators identified by the claimed method are useful for

[illegible]

RESULT 11
ID W35946 standard; Protein; 604 AA.
AC W35946;
DI 11-MAR-1998 (first entry)
DE Human neurin-1.
KW Netrin-1; neuron; growth; differentiation; morphology;
KW Netrin-1; disease; diagnosis; therapy; drug screening; human.
OS Homo sapiens.
PN W09740064-A1.

PR 19-APR-1996; US-6355137.
PA (EXEL-) EXELIXIS PHARM INC.
PI (REGC.) UNIV CALIFORNIA
PI Kennedy T, Leonardo D, Serafini T, Shyian A, Swimmer C,
PI Tessier-Lavigne M, Zhang F;
PI WPI: 97-535773/49.
DR N-PSDB: T97129.
DR
PT Human netrin-1 protein and related nucleic acids - useful in
PT modulating neuron growth and screening for compounds for diagnosis

PT or treatment of diseases associated with undesirable growth
PS Claim 1; Page 13-15; 22pp: English.
CC This protein comprises human netrin-1 (see w35946), a protein
CC which is involved in neural axon guidance, and which is especially
CC useful in modulating neural axon outgrowth. Its amino acid
CC sequence was deduced from a cDNA clone (see T97129) isolated from a
CC human foetal brain cDNA library. Neuron growth, differentiation or
CC morphology can be altered by contact with netrin (claimed).
CC Isolated netrin-1 can also be used to screen chemical libraries
CC for candidate drugs suitable for diagnosis or treatment of diseases
CC associated with undesirable neural cell growth, by comparing
CC binding to a netrin binding target with and without the presence of
CC a prospective agent. Agents that modulate the interaction may be
CC useful as pharmaceutical lead compounds (claimed).
SQ Sequence 504 AA;

Query Match 5.4%; Score 121; DB 28; Length 604;
Best Local Similarity 24.8%; Pred. No. 8,16e-02;
Matches 29; Conservative 30; Mismatches 49; Indels 9; Gaps

[illegible]

```

FT      /note= "unidentified amino acids"
FT
FT      misc_difference 300..301
FT      /note= "unidentified amino acids"
FT

```

FT		/note=	"unidentified amino acids"
FT			
FT	misc_difference	311	
FT		/note=	"unidentified amino acid"
FT			
FT	misc_difference	389	
FT		/note=	"unidentified amino acid"
FT			
FT	misc_difference	393..394	
FT		/note=	"unidentified amino acids"

FT	misc_difference	487	
FT			/note= "unidentified amino acid"

PN W09513367-A1.
 PD 18-MAY-1995.
 PF 08-NOV-1994; U12913.
 PR 12-NOV-1993; US-152019.
 PA (REGC) UNIV CALIFORNIA.
 PA (UYCO) UNIV COLUMBIA NEW YORK.
 PI Dodd J, Jessell T, Kennedy T, Placzek M, Serafini T;
 PI Tessier-Lavigne M;
 DR WPI: 95-194086/25.
 DR N-PSDB: Q92388.
 PT Neural axon out-growth modulators derived from EGF-like repeats of
 PT netrin 1 or netrin 2 - comprise peptide(s) capable of selectively
 PT increasing spinal axon out-growth or directing axon orientation
 PS Claim 1; Page 51-52; 58pp; English.
 CC Chick p75 and p78 sequences (given in R74186-87, respectively)
 CC were used to identify conserved amino acid regions, which were
 CC then used to design degenerate primers for the amplification
 CC of a fragment of the mouse p78 cDNA (Q92368). This cDNA can
 CC be expressed in host cells for recombinant p78 prodn, or used
 CC to breed transgenic animals, or for gene therapy.
 SQ Sequence 529 AA;
 Query Match 4.7%; Score 105; DB 14; Length 529;
 Best Local Similarity 23.5%; Pred. NO. 1.70e+00;
 Matches 24; Conservative 25; Mismatches 45; Indels 8; Gaps 7;
 Db 398 sxveexdcysckaskgkikmmkkyckrdyavqihlkadkagdw-wkftvniis-- 444
 Qy 163 PATEAPKVCBA-CNKNDNDNDMETLCKNDFAKTKV-KEITYINDYKILLETSKT 220
 Db 445 vykgtgrirrgdgs-lwrsrdiackepkikplk-ylllg 484
 Qy 221 IYKLVGSERLKKSVMWLD-SLQCTCEMNDNAPLYVMG 261
 RESULT 13
 ID W26642 standard; Protein; 732 AA.
 AC W26642;
 DT 11-FEB-1998 (first entry)
 DE Human RECK cancer-inhibiting protein.
 KW RECK; reversion-inducing cysteine rich protein with Kazal motif;
 KW human; cancer; gene therapy; diagnosis.
 OS Homo sapiens.
 FH Key Location/Qualifiers
 FT Domain 396..415
 FT /label= Kazal_domain
 PN W09724439-A1.
 PD 10-JUL-1997.
 PF 24-DEC-1996; U20812.
 PR 27-DEC-1995; JP-340469.
 PA (AMGE-) AMGEN INC.
 PA (KITA/) KITAYAMA H.
 PA (NODA/) NODA M.
 PA (TAKA/) TAKAHASHI C.
 PA (SANY) SANKYO CO LTD.
 PI Kitayama H, Noda M, Takahashi C;
 DR WPI: 97-363675/33.
 DR N-PSDB: T90508.
 PT RECK gene and corresponding protein sequences - enables reversion of
 PT cancer cells
 PS Claim 7; Page 42-44; 53pp; English.
 CC This protein sequence comprises the human RECK protein (reversion-
 CC inducing cysteine rich protein with Kazal motif), which is capable
 CC of transforming malignant cancer cells with an activated ras gene
 CC into normal cells (reversion activity). Its amino acid sequence
 CC was deduced from a cDNA clone (see T90508) obtained from human
 CC fibroblast MRC-5 (ATCC CCL 171) cells. A claimed method for
 CC treatment of cancer comprises contacting the cancer cells with a
 CC RECK polypeptide. RECK may also be expressed using gene therapy
 CC methods for in vivo treatment of cancer.
 SQ Sequence 732 AA;
 Query Match 4.5%; Score 100; DB 25; Length 732;

Best Local Similarity 30.5%; Pred. NO. 4.26e+00;
 Matches 25; Conservative 15; Mismatches 36; Indels 6; Gaps 5;
 Db 158 ipvldikkcpemwkaiaclqikpchkshksrsil-cksdceveillkkcgd-qnkfpdht 215
 Qy 80 IP-LVMKQCHPDT-KKFLCSLFAFVCLDLDDETQPCSLCVQVKRCAPVMSAFGPPWP 137
 Db 216 aesicellsptddkncipldt 237
 Qy 138 DMLECDRFPQDNDL--CIPLAS 157
 RESULT 14
 ID W55368 standard; Protein; 804 AA.
 AC W55368;
 DT 17-JUN-1998 (first entry)
 DE H. pylori ORF 09apl1406orf2 protein.
 KW Cytoplasmic; vaccine; prevention; treatment; infection; envelope;
 KW identification; binding compound; bacteria; life cycle; activator;
 KW inhibitor; duodenal ulcer disease; chronic gastritis; diagnosis.
 OS Helicobacter pylori.
 PN W09737044-A1.
 PD 09-OCT-1997.
 PF 27-MAR-1997; U05223.
 PR 06-DEC-1996; US-761318.
 PR 29-MAR-1996; US-625811.
 PR 02-APR-1996; US-758731.
 PR 25-OCT-1996; US-736905.
 PR 28-OCT-1996; US-738859.
 PA (ASTR) ASTRA AB.
 PI Alm RA, Smith D;
 DR WPI: 97-503122/46.
 DR N-PSDB: V24777.
 PT Helicobacter pylori nucleic acid sequences and encoded
 PT polypeptide(s) - useful in vaccines to treat or prevent H. pylori
 PT infection and for diagnosis of H. pylori infection
 PS Claim 14; Page 587-589; 1145pp; English.
 CC This sequence is a H. pylori protein of unspecified function.
 CC The protein may be used in a vaccine to prevent or treat H. pylori
 CC infection or to identify H. pylori polypeptide binding compounds. The
 CC useful as potential H. pylori life cycle activators or inhibitors. The
 CC DNA and probes derived from it may be used for the identification of
 CC H. pylori in a sample and the diagnosis of H. pylori infection. Nucleic
 CC acid sequences complementary to the DNA act as antisense sequences and
 CC can be used to prevent the translation of H. pylori mRNA. Antibodies
 CC against the protein can be used in immunoassays to evaluate the abundance
 CC and distribution of H. pylori-specific antigens. The genomic sequence of
 CC H. pylori (ATCC 35679) was determined from overlapping contigs generated
 CC by mechanically shearing the bacterial DNA. The sequences were analysed
 CC for ORF of at least 180 nucleotides, and the predicted coding regions
 CC defined by computer evaluation. To identify likely H. pylori antigens for
 CC vaccine development, the amino acid sequences predicted from various ORF
 CC were analysed for significant homology to other known or exported
 CC membrane proteins. Having identified and determined the sequences of
 CC interest, particular regions can be isolated from H. pylori by PCR
 CC amplification for recombinant polypeptide production, e.g. in E. coli
 CC hosts.
 SQ Sequence 804 AA;
 Query Match 4.4%; Score 99; DB 29; Length 804;
 Best Local Similarity 31.5%; Pred. NO. 5.11e+00;
 Matches 17; Conservative 16; Mismatches 19; Indels 2; Gaps 2;
 Db 78 vidsipkeftplenaftnysmrqgqfihikpkmvrrltlfsfdrydkkai 131
 Qy 185 IMETLCKNDFA-LKIKVKEITINRDTKILLETSKTIYKLVGS-ERDLKKS 236
 RESULT 15
 ID W26610 standard; Protein; 868 AA.
 AC W26610;
 DT 27-JAN-1998 (first entry)
 DE Rat muscle-specific kinase (Musk).

Search completed: Tue Aug 4 10:32:08 1998
Job time : 48 secs.

THIS PAGE BLANK (USPTO)

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

```
>US-08-848-439-2
Title:
Description: (1-295) from US08848439.pep
Perfect Score: 2234
Sequence: 1 MLOGPGSLLFLFASHCCGL.....WQKQREFFKRISIRKLOC 295
```

Post-processing: Minimum Match 0%
Listing first 45 summaries

Statistics: Mean 44.927; variance 86.902; scale 0.517

SUMMARIES

Result	No.	Score	Query	Match	Length	DB	ID	Description	Pred. No.
1	386	17.3	641	2	A45054			probable intercellula	1.70e-52
2	355	15.9	694	2	S78444			dfz2 protein - fruit	2.44e-46
3	355	15.9	694	2	S71786			dfz2 protein - fruit	2.44e-46
4	332	14.9	415	2	S15709			hypothetical protein	8.11e-42
5	332	14.9	581	2	S03340			gene frizzled protein	8.11e-42
6	178	8.0	562	2	B6101			collagen alpha 1(XVII	3.36e-13
7	148	6.6	581	2	B54655			netrin-2 precursor	3.23e-08
8	124	5.6	606	2	A54865			netrin-1 precursor	1.67e-04
9	122	5.5	793	2	JC5339			Smoothed protein -	3.30e-04
10	105	4.7	500	2	D31579			cytochrome P450 2D4	8.40e-02
11	105	4.7	500	2	I52313			cytochrome P450 2D4 h	8.40e-02
12	101	4.5	1735	2	S28812			DNA-directed RNA poly	2.87e-01
13	99	4.4	322	2	H64825			hypothetical protein	5.25e-01
14	98	4.4	391	2	E64474			hypothetical protein	7.08e-01
15	99	4.4	724	2	B32571			ribosomal protein S6	5.25e-01
16	98	4.4	1032	1	H7BPT4			baseplate protein gp7	7.08e-01
17	96	4.3	520	2	J50291			intermediate filament	1.28e+00
18	96	4.3	735	2	I51901			ribosomal protein S6	1.28e+00
19	97	4.3	15281	2	S41309			cyclosporin synthetas	9.52e-01
20	94	4.2	771	2	E64594			hypothetical protein	2.28e+00
21	93	4.2	735	2	A53300			ribosomal protein S6	3.04e+00
22	92	4.1	203	2	S70254			conserved hypothetica	4.05e+00
23	92	4.1	608	2	B70188			hypothetical protein	4.05e+00

```
ACCESSIONS S78444
REFERENCE S78444
#authors Bhanot, P.; Wang, Y.; Nathans, J.
#submission submitted to the EMBL Data Library, July 1996
#accession S78444
##status preliminary
##molecule_type DNA
##residues 1-694 #label BHA
##cross-references EMBL:U65589
SUMMARY #length 694 #molecular-weight 75437 #checksum 3113

Query Match 15.9%; Score 355; DB 2; Length 694;
Best Local Similarity 36.2%; Pred. No. 2.44e-46;
Matches 47; Conservative 32; Mismatches 45; Indels 6; Gaps 6;

Db 51 YGVPAIP-KDPNLRCEIIPMCRGIGYNTSFPNEMNHETODEAGLEVHQFW-PLVEIK 108
QY 28 FGQPDFSYKRSNCKPIPANLQCHGIEYQNMRLPNLLGHETMKEV-LEQAGAWIPLVMKQ 86
Db 109 CSPDLKFFLCSMYTPICLDYHKPLPVCRSVCERARSGCAPIMQOYSFEWPERMACEHLP 168
QY 87 CHPDTKKFLCSLFPAPVCLDLDLDTIQCHSLCVQVKDRCAPVMSAFGFPWPDMLECDREFP 146
Db 169 LHGDPDNLNM 178
QY 147 -Q-D-NDLCI 153

RESULT 3
ENTRY S71786 #type complete
TITLE dfz2 protein - fruit fly (Drosophila melanogaster)
ORGANISM #formal_name Drosophila melanogaster
DATE 12-Feb-1998 #sequence_revision 13-Mar-1998 #text_change
13-Mar-1998
ACCESSIONS S71786
REFERENCE S71786
#authors Bhanot, P.; Brink, M.; Harryman Samos, C.; Hsieh, J.C.; Wang,
Y.; Macke, J.P.; Andrew, D.; Nathans, J.; Nusse, R.
#journal Nature (1996) 382:225-230
#title A new member of the frizzled family from Drosophila functions
as a wingless receptor.
#accession S71786
##status preliminary; nucleic acid sequence not shown
##molecule_type DNA
##residues 1-694 #label BHA
##cross-references EMBL:U65589
SUMMARY #length 694 #molecular-weight 75423 #checksum 3095

Query Match 15.9%; Score 355; DB 2; Length 694;
Best Local Similarity 36.2%; Pred. No. 2.44e-46;
Matches 47; Conservative 32; Mismatches 45; Indels 6; Gaps 6;

Db 51 YGVPAIP-KDPNLRCEIIPMCRGIGYNTSFPNEMNHETODEAGLEVHQFW-PLVEIK 108
QY 28 FGQPDFSYKRSNCKPIPANLQCHGIEYQNMRLPNLLGHETMKEV-LEQAGAWIPLVMKQ 86
Db 109 CSPDLKFFLCSMYTPICLDYHKPLPVCRSVCERARSGCAPIMQOYSFEWPERMACEHLP 168
QY 87 CHPDTKKFLCSLFPAPVCLDLDLDTIQCHSLCVQVKDRCAPVMSAFGFPWPDMLECDREFP 146
Db 169 LHGDPDNLNM 178
QY 147 -Q-D-NDLCI 153

RESULT 4
ENTRY S15709 #type complete
TITLE hypothetical protein - fruit fly (Drosophila melanogaster)
ORGANISM #formal_name Drosophila melanogaster
DATE 21-Nov-1993 #sequence_revision 10-Nov-1995 #text_change
16-Feb-1997
ACCESSIONS S15709
REFERENCE S15708

#authors Adler, P.N.; Vinson, C.; Park, W.J.; Conover, S.; Klein, L.
#journal Genetics (1990) 126:401-416
#title Molecular structure of frizzled, a Drosophila tissue polarity
gene.
#cross-references MUID:91060073
#accession S15709
##status preliminary
##molecule_type DNA
##residues 1-581 #label ADL
##cross-references EMBL:X54648
GENETICS
#gene FlyBase:frz
#cross-references FlyBase:FBgn0001085
#introns 224/3; 264/3; 329/3; 405/1
KEYWORDS alternative splicing; transmembrane protein
FEATURE 1-26
27-581
#domain signal sequence #status predicted #label SIGV
#product gene frizzled protein #status predicted #label
MAT
SUMMARY #length 581 #molecular-weight 64846 #checksum 358

Query Match 14.9%; Score 332; DB 2; Length 581;
Best Local Similarity 37.5%; Pred. No. 8.11e-42;
Matches 45; Conservative 22; Mismatches 47; Indels 6; Gaps 6;
```

```
#authors Adler, P.N.; Vinson, C.; Park, W.J.; Conover, S.; Klein, L.
#journal Genetics (1990) 126:401-416
#title Molecular structure of frizzled, a Drosophila tissue polarity
gene.
#cross-references MUID:91060073
#accession S15709
##status preliminary
##molecule_type DNA
##residues 1-415 #label ADL
##cross-references EMBL:X54648
GENETICS
#gene FlyBase:frz
#cross-references FlyBase:FBgn0001085
#introns 224/3; 264/3; 329/3; 405/1
SUMMARY #length 415 #molecular-weight 46117 #checksum 4401

Query Match 14.9%; Score 332; DB 2; Length 415;
Best Local Similarity 37.5%; Pred. No. 8.11e-42;
Matches 45; Conservative 22; Mismatches 47; Indels 6; Gaps 6;

Db 50 HNRCEPI-T-ISICKNIPYNTIMPNLIGHTKQEEAGLEVHQFAPLVKIGSCDDLQFLC 107
QY 37 RSNCAPIPANLQCHGIEYQNMRLPNLLGHETMKEVLEQAGAWIPLVMKQCHPDTKKFLC 96
Db 108 SLYPVC-TILERPPIPCRSLSARV-CEKLMKTYNFWNPENLECSKFPVHGGEDLCVA 165
QY 97 SLFAPVCLDLDLDTIQCHSLCVQVKDRCAPVMSAFGFPWPDMLECDREFP-QD-NDLCIP 154

RESULT 5
ENTRY S03540 #type complete
TITLE gene frizzled protein precursor - fruit fly (Drosophila
melanogaster)
ORGANISM #formal_name Drosophila melanogaster
DATE 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change
23-Aug-1997
ACCESSIONS S03540; S15708
REFERENCE S03540
#authors Vinson, C.R.; Conover, S.; Adler, P.N.
#journal Nature (1989) 338:263-264
#title A Drosophila tissue polarity locus encodes a protein
containing seven potential transmembrane domains.
#cross-references MUID:89159415
#accession S03540
##status not compared with conceptual translation
##molecule_type mRNA
##residues 1-581 #label VIN
REFERENCE S15708
#authors Adler, P.N.; Vinson, C.; Park, W.J.; Conover, S.; Klein, L.
#journal Genetics (1990) 126:401-416
#title Molecular structure of frizzled, a Drosophila tissue polarity
gene.
#cross-references MUID:91060073
#accession S15708
##status preliminary
##molecule_type DNA
##residues 1-581 #label ADL
##cross-references EMBL:X54648
GENETICS
#gene FlyBase:frz
#cross-references FlyBase:FBgn0001085
#introns 224/3; 264/3; 329/3; 405/1
KEYWORDS alternative splicing; transmembrane protein
FEATURE 1-26
27-581
#domain signal sequence #status predicted #label SIGV
#product gene frizzled protein #status predicted #label
MAT
SUMMARY #length 581 #molecular-weight 64846 #checksum 358

Query Match 14.9%; Score 332; DB 2; Length 581;
Best Local Similarity 37.5%; Pred. No. 8.11e-42;
Matches 45; Conservative 22; Mismatches 47; Indels 6; Gaps 6;
```

```

Db 50 HNRCEPI-T-ISICKNIPYNMTIMPNLIGHTKQEBAGLEVHQFAPLVYKIGSDDDLQFLC 107
::: ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : |||
Qy 37 RSNCKPIPANLQCHGIEYQNMRLPNLLGHETMKVELEQAGAWIPLVNKQCHPDTKKFLC 96
::: ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : |||
Db 108 SLYVPVC-TILERTPPCRSICESARV-CEKLMKTYNENWPNLECSKFPVHGGECLVA 165
::: ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : |||
Qy 97 SLFAPVCLDDLDLDTIQPCHSLCVQVKDRCAPVMVSAFGFPWMDLECDRFP-QD-NDLCIP 154
::: ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : |||

RESULT 6
ENTRY B56101 #type fragment
TITLE collagen alpha 1(XVII) chain precursor long form - mouse
ORGANISM (fragment)
#formal_name Mus musculus #common_name house mouse
DATE 03-Oct-1995 #sequence_revision 03-Oct-1995 #text_change
10-Sep-1997
ACCESSIONS B56101
REFERENCE A56101
#authors Rehn, M.; Pihlajaniemi, T.
#journal J. Biol. Chem. (1995) 270:4705-4711
#title Identification of three N-terminal ends of type XVIII
collagen chains and tissue-specific differences in the
expression of the corresponding transcripts. The longest
form contains a novel motif homologous to rat and
Drosophila frizzled proteins.
#accession B56101
##status preliminary
##molecule_type mRNA
##residues 1-562 #label REH
##cross-references GB:U11637; NID:9618429; PID:9618430
GENETICS
#gene COL18A1
SUMMARY #length 562 #checksum 115

Query Match 8.0%; Score 178; DB 2; Length 562;
Best Local Similarity 34.1%; Pred. No. 3.36e-13;
Matches 30; Conservative 13; Mismatches 45; Indels 0; Gaps 0;

Db 368 SRCLPPLPTLCRLGIGHFPLNHLHTDSVEATVQAWGRFLHTNCHPFLAWFCL 427
||| : : : ||| : : : ||| : : : ||| : : : ||| : : : |||
Qy 38 SNCRPIPANLQCHGIEYQNMRLPNLLGHETMKVELEQAGAWIPLVNKQCHPDTKKFLCS 97
||| : : : ||| : : : ||| : : : ||| : : : ||| : : : |||

Db 428 LLAPSCGPPPPPLPPCRQCFCEALDEC 455
||| : : : ||| : : : ||| : : : ||| : : : |||
Qy 98 LFAPVCLDDLDLDTIQPCHSLCVQVKDRC 125
||| : : : ||| : : : ||| : : : ||| : : : |||

RESULT 7
ENTRY B54665 #type fragment
TITLE netrin-2 precursor - chicken (fragment)
ORGANISM #formal_name Gallus gallus #common_name chicken
DATE 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change
10-Sep-1997
ACCESSIONS B54665
REFERENCE A54665
#authors Serafini, T.; Kennedy, T.E.; Galko, M.J.; Mirzayan, C.;
Jessel, T.M.; Tessier-Lavigne, M.
#journal Cell (1994) 78:409-424
#title The netrins define a family of axon outgrowth-promoting
proteins homologous to Caenorhabditis elegans UNC-6.
#accession B54665
##status preliminary
##molecule_type mRNA
##residues 1-581 #label SER
##cross-references GB:L34550; NID:9529420; PID:9529421
SUMMARY #length 581 #checksum 7731

Query Match 6.6%; Score 148; DB 2; Length 581;
Best Local Similarity 27.0%; Pred. No. 3.23e-08;
Matches 30; Conservative 28; Mismatches 46; Indels 7; Gaps 6;

Db 431 IPAINPTSLVSTSTE-APADCDSCYCKPAKGNKYKINMKYCKRDYVQVNVNILEMETVANWAK 489
||| : : : ||| : : : ||| : : : ||| : : : ||| : : : |||

```

```

Qy 153 IPLASSDHLLPATEAPKVCEA-CKNKNDNDDNDIMETLCNKDNDFALKTKVKEITYINRDTK 211
Db 490 FTNILS--VYKCRDERVKR-GDNFLHILKDLSCCKPKI-QISKKYLVMG 536
::: ||| : : : ||| : : : ||| : : : ||| : : : |||
Qy 212 IILETKSKTIYKLVNGVSEDLKKSVLWK-DSLQCTCEENNDINAPYLVMG 261
::: ||| : : : ||| : : : ||| : : : ||| : : : |||

RESULT 8
ENTRY A54665 #type complete
TITLE netrin-1 precursor - chicken
ORGANISM #formal_name Gallus gallus #common_name chicken
DATE 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change
10-Sep-1997
ACCESSIONS A54665
REFERENCE A54665
#authors Serafini, T.; Kennedy, T.E.; Galko, M.J.; Mirzayan, C.;
Jessel, T.M.; Tessier-Lavigne, M.
#journal Cell (1994) 78:409-424
#title The netrins define a family of axon outgrowth-promoting
proteins homologous to Caenorhabditis elegans UNC-6.
#accession A54665
##status preliminary
##molecule_type mRNA
##residues 1-606 #label SER
##cross-references GB:L34549; NID:9529418; PID:9529419
SUMMARY #length 606 #molecular-weight 68126 #checksum 5400

Query Match 5.6%; Score 124; DB 2; Length 606;
Best Local Similarity 26.1%; Pred. No. 1.67e-04;
Matches 31; Conservative 34; Mismatches 41; Indels 13; Gaps 11;

Db 456 IPAAPPTTASSTEEPAD-CDSYCKASKGKLKINMKYCKRDYAVQIHLKAENADM-W 513
||| : : : ||| : : : ||| : : : ||| : : : ||| : : : |||
Qy 153 IPLASSDHLLPATEAPKVCEA-CKNKNDNDDNDIMETLCNKDNDFALKTKV-KEITYINRDT 210
||| : : : ||| : : : ||| : : : ||| : : : ||| : : : |||

Db 514 KFTVNIIS--VYK-OG-SNR-LRRGDOTLVHAKDIACKCPKVPKMK-YLLLGSTEDS 566
||| : : : ||| : : : ||| : : : ||| : : : ||| : : : |||
Qy 211 KIILETKSKTIYKLVNGVSEDLKKS--VLWLKD-SLQCTCEENNDINAPYLVMGQKGG 266
||| : : : ||| : : : ||| : : : ||| : : : ||| : : : |||

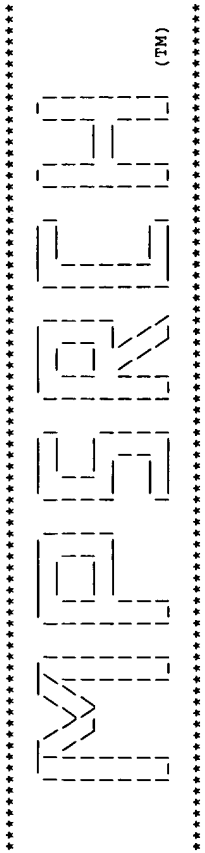
RESULT 9
ENTRY JC5539 #type complete
TITLE Smoothed protein - mouse
ORGANISM #formal_name Mus musculus #common_name house mouse
DATE 02-Sep-1997 #sequence_revision 05-Sep-1997 #text_change
05-Sep-1997
ACCESSIONS JC5539; PC4476
REFERENCE JC5539
#authors Akiyama, H.; Shigeno, C.; Hiraki, Y.; Shukunami, C.; Kohnno,
H.; Akagi, M.; Konishi, J.; Nakamura, T.
#journal Biochem. Biophys. Res. Commun. (1997) 235:142-147
#title Cloning of a mouse Smoothed cDNA and expression patterns of
hedgehog signalling molecules during chondrogenesis and
cartilage differentiation in clonal mouse EC cells, ADPC5.
#accession JC5539
##molecule_type mRNA
##residues 1-793 #label AKI
#accession PC4476
##molecule_type protein
##residues 528-533:539-545;600-605 #label AK2
##experimental_source ADPC5 cell
COMMENT This protein is used in the conserved targets in Hedgehog
signalling pathway, together with Patched and Gli. These protein
are responsible for the skeletal abnormalities in Gorlin and
Greig syndromes.
FEATURE
1-32 #domain signal sequence #status predicted #label SIG
SUMMARY #length 793 #molecular-weight 87299 #checksum 8609

Query Match 5.5%; Score 122; DB 2; Length 793;
Best Local Similarity 26.7%; Pred. No. 3.30e-04;
Matches 31; Conservative 22; Mismatches 54; Indels 9; Gaps 7;

```


GENETICS

THIS PAGE BLANK (USPTO)



Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPsrch_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Tue Aug 4 10:33:44 1998; Maspar time 10.67 Seconds
693.479 Million cell updates/sec

Tabular output not generated.

Title: >US-08-848-439-2
Description: (1-295) from US08848439.pep
Perfect Score: 2234
Sequence: 1 MLQGGSGLLLLFLASHCLG.....WKGQREKFRKRSIRKLQC 295

Scoring table: PAM 150
Gap 11

Searched: 69111 seqs, 25083644 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: swiss-prot35
1:swiss1

Statistics: Mean 46.742; Variance 71.562; scale 0.653

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	332	14.9	581	1	FRIZ_DROME FRIZZLED PROTEIN PRECU	9.66e-53
2	148	6.6	581	1	NETR2_CHICK NETRIN-2 PRECURSOR (FR	8.14e-11
3	124	5.6	606	1	NETR1_CHICK NETRIN-1 PRECURSOR.	3.06e-06
4	109	4.9	465	1	YQ18_CAEEL HYPOTHETICAL 53.8 KD P	1.32e-03
5	105	4.7	500	1	CPD4_RAT CYTOCHROME P450 IID4 (6.13e-03
6	105	4.7	500	1	CPD1_RAT CYTOCHROME P450 IID18 (6.13e-03
7	101	4.5	1741	1	PCP1_GIALA DNA-DIRECTED RNA POLYM	2.73e-02
8	99	4.4	724	1	RG61_MOUSE RIBOSOMAL PROTEIN S6 K	5.67e-02
9	98	4.4	1032	1	VG07_BPT4 BASEPLATE STRUCTURAL P	8.13e-02
10	96	4.3	520	1	ION3_CARAU INTERMEDIATE FILAMENT	1.66e-01
11	94	4.2	531	1	UD13_RAT UDP-GLUCURONOSYLTRANSF	3.35e-01
12	92	4.1	1562	1	YM81_YEAST HYPOTHETICAL 180.2 KD	6.69e-01
13	90	4.0	216	1	GYRB_ACIS3 DNA GYRASE SUBUNIT B (1.31e+00
14	89	4.0	262	1	VIRF_SHIDY VIRULENCE REGULON TRAN	1.83e+00
15	90	4.0	263	1	KSGA_MFCPN DIMETHYLADENOSINE TRAN	1.31e+00
16	90	4.0	561	1	VATA_MAIZE VACUOLAR ATP SYNTHASE	1.31e+00
17	89	4.0	752	1	KSGA_CHICK RIBOSOMAL PROTEIN S6 K	1.83e+00
18	88	3.9	127	1	YB09_YEAST HYPOTHETICAL 14.8 KD P	2.55e+00
19	88	3.9	182	1	YLP2_CAEEL HYPOTHETICAL 20.1 KD P	2.55e+00
20	88	3.9	443	1	AMZA_ORISA ALPHA-AMYLASE ISOZYME	2.55e+00
21	88	3.9	445	1	AMC2_ORISA ALPHA-AMYLASE ISOZYME	2.55e+00
22	87	3.9	446	1	YHE2_PSEAE HYPOTHETICAL 50.0 KD P	3.53e+00
23	87	3.9	510	1	NOA1_HUMAN ONCONEURAL VENTRAL ANT	3.53e+00

24	87	3.9	580	1	VATA_HORVU VACUOLAR ATP SYNTHASE	3.53e+00
25	88	3.9	633	1	KS62_MOUSE RIBOSOMAL PROTEIN S6 K	2.55e+00
26	88	3.9	740	1	KS62_HUMAN RIBOSOMAL PROTEIN S6 K	2.55e+00
27	88	3.9	848	1	NFM_MOUSE NEUROFILAMENT TRIPLET	2.55e+00
28	88	3.9	1102	1	YK64_CAEEL HYPOTHETICAL 124.8 KD	2.55e+00
29	87	3.9	1201	1	COPA_YEAST COTOMER ALPHA SUBUNIT	3.53e+00
30	87	3.9	1752	1	DESP_HUMAN DESMOPLAKIN I AND II (3.53e+00
31	87	3.9	2386	1	FINC_HUMAN FIBRONECTIN PRECURSOR	3.53e+00
32	87	3.9	2749	1	IP3R_MOUSE INOSITOL 1,4,5-TRISPHO	3.53e+00
33	87	3.9	2749	1	IP3R_RAT INOSITOL 1,4,5-TRISPHO	3.53e+00
34	86	3.8	327	1	VMP_CAMVD MOVEMENT PROTEIN (CELL	4.88e+00
35	86	3.8	450	1	DCOR_CHICK ORNITHINE DECARBOXYLAS	4.88e+00
36	86	3.8	461	1	DCOR_BOVIN ORNITHINE DECARBOXYLAS	4.88e+00
37	85	3.8	495	1	E1BL_ADE02 E1B PROTEIN, LARGE T-A	6.72e+00
38	86	3.8	556	1	FTHS_STRUM FORMATE--TETRAHYDROFOL	4.88e+00
39	86	3.8	620	1	Y870_METJA HYPOTHETICAL PROTEIN M	4.88e+00
40	86	3.8	725	1	VACB_MYCGE VACB PROTEIN HOMOLOG	4.88e+00
41	85	3.8	798	1	VP16_YEAST VACUOLAR PROTEIN SORTI	6.72e+00
42	85	3.8	1701	1	MSPI_PLAFF MEROZOITE SURFACE PROT	6.72e+00
43	85	3.8	1726	1	MSPI_PLAFC MEROZOITE SURFACE PROT	6.72e+00
44	85	3.8	1726	1	MSPI_PLAFC MEROZOITE SURFACE PROT	6.72e+00
45	85	3.8	2688	1	2EPI_MOUSE ZINC FINGER PROTEIN 40	6.72e+00

ALIGNMENTS

RESULT 1

ID FRIZ_DROME STANDARD; PRT; 581 AA.

AC P18537; 01-NOV-1990 (REL. 16, CREATED)

DT 01-NOV-1990 (REL. 16, LAST SEQUENCE UPDATE)

DT 01-MAY-1992 (REL. 22, LAST ANNOTATION UPDATE)

DE FRIZZLED PROTEIN PRECURSOR.

GN FZ.

OS DROSOPHILA MELANOGASTER (FRUIT FLY).

OC EUKARYOTA; METAZOA; ARTHROPODA; INSECTA; DIPTERA.

RN [1]

RP SEQUENCE FROM N.A. (CLASS I).

RX MEDLINE; 89159415.

RA VINSON C.R., CONOVER S., ADLER P.N.;

RL NATURE 338:263-264(1989).

[2]

RP SEQUENCE FROM N.A. (CLASSES I AND II).

RX MEDLINE; 91060073.

RA ADLER P.N., VINSON C., PARK W.J., CONOVER S., KLEIN L.;

RL GENETICS 126:401-416(1990).

CC -!- FUNCTION: REQUIRED TO COORDINATE THE CYTOSKELETONS OF EPIDERMAL CELLS TO PRODUCE A PARALLEL ARRAY OF CUTICULAR HAIRS AND BRISTLES. FZ IS REQUIRED FOR THE TRANSMISSION OF POLARITY INFORMATION, AND TRANSDUCTION.

CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.

CC -!- ALTERNATIVE PRODUCTS: CLASS I AND CLASS II PROTEINS DIFFER AT THE C-TERMINAL FROM AA 406 DUE TO USE OF ALTERNATIVE 3' EXONS.

DR EMBL; X54648; G804979; JOINED.

DR EMBL; X54649; G804979; JOINED.

DR EMBL; X54650; G804979; JOINED.

DR EMBL; X54651; G804979; JOINED.

DR EMBL; X54648; G804980; JOINED.

DR EMBL; X54649; G804980; JOINED.

DR EMBL; X54650; G804980; JOINED.

DR EMBL; X54652; G804980; JOINED.

DR EMBL; X54646; G7981; .

DR EMBL; X54647; G7983; ALT_SEQ.

DR PIR; S03540; S03540.

DR PIR; S15708; S15708.

DR FLYBASE; FBgn0001085; fz.

KW TRANSMEMBRANE; SIGNAL; ALTERNATIVE SPLICING; DEVELOPMENTAL PROTEIN.

FT SIGNAL 1 26

FT CHAIN 27 581

FT FRIZZLED PROTEIN.

FT TRANSMEM 248 270

FT TRANSMEM 281 303

FT TRANSMEM 336 368

FT TRANSMEM 381 401

RESULT	3	STANDARD;	PRT;	606 AA.
ID	NET1_CHICK			
AC	Q50922;			
DT	01-NOV-1997 (REL. 35, CREATED)			
DT	01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)			
DT	01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)			
DE	NETRIN-1 PRECURSOR.			
OS	GALLUS GALLUS (CHICKEN).			
OS	EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; AVES; NEOGNATHAE.			
OC	GALLIFORMES.			
NC	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN-WHITE LEGHORN; TISSUE-EMBRYONIC BRAIN;			
RC	MEDLINE; 94340732.			
RA	SERAFINI T., KENNEDY T.E., GALKO M.J., MIRZAYAN C., JESSELL T.M.,			
RA	TESSIER-LAVIGNE M.;			
RL	CELL 78:409-428 (1994).			
CC	-1- FUNCTION: NETRINS CONTROL GUIDANCE OF CNS COMMISSURAL AXONS AND PERIPHERAL MOTOR AXONS.			
CC	-1- SUBCELLULAR LOCATION: EXTRACELLULAR.			
CC	-1- SIMILARITY: CONTAINS ONE LAMININ N-TERMINAL DOMAIN (DOMAIN VI).			
CC	-1- SIMILARITY: CONTAINS 3 LAMININ EGF-LIKE DOMAINS.			
CC	-1- SIMILARITY: CONTAINS ONE C345C DOMAIN.			
CC	-1- SIMILARITY: TO CAENORHABDITIS ELEGANS UNC-6 PROTEIN.			
DR	EMBL; L34549; G529419; "			
DR	PROSITE; PS00022; EGF.1; 2.			
DR	PROSITE; PS01248; LAMININ TYPE EGF.3.			
KW	GLYCOPROTEIN; EXTRACELLULAR MATRIX; SIGNAL; LAMININ EGF-LIKE DOMAIN;			
KW	REPEAT.			
FT	SIGNAL	1	25	POTENTIAL.
FT	CHAIN	26	606	NETRIN-1.
FT	DOMAIN	26	286	3 X LAMININ N-TERMINAL (DOMAIN VI).
FT	DOMAIN	287	455	3 X LAMININ EGF-LIKE REPEATS (LAMININ DOMAIN V-LIKE).
FT	DOMAIN	287	342	LAMININ EGF-LIKE 1.
FT	DOMAIN	343	405	LAMININ EGF-LIKE 2.
FT	DOMAIN	406	455	LAMININ EGF-LIKE 3.
FT	DOMAIN	456	606	C345C (DOMAIN C).
FT	SITE	532	534	CELL ATTACHMENT SITE (POTENTIAL).
FT	DISULFID	121	154	BY SIMILARITY.
FT	DISULFID	287	296	BY SIMILARITY.
FT	DISULFID	289	306	BY SIMILARITY.
FT	DISULFID	308	317	BY SIMILARITY.
FT	DISULFID	320	340	BY SIMILARITY.
FT	DISULFID	343	352	BY SIMILARITY.
FT	DISULFID	345	370	BY SIMILARITY.
FT	DISULFID	373	382	BY SIMILARITY.
FT	DISULFID	385	403	BY SIMILARITY.
FT	DISULFID	406	418	BY SIMILARITY.
FT	DISULFID	408	425	BY SIMILARITY.
FT	DISULFID	427	436	BY SIMILARITY.
FT	DISULFID	439	453	BY SIMILARITY.

FT	DISULFID	478	546	BY SIMILARITY.
FT	DISULFID	493	603	BY SIMILARITY.
FT	CARBOHYD	97	97	POTENTIAL.
FT	CARBOHYD	118	118	POTENTIAL.
FT	CARBOHYD	133	133	POTENTIAL.
FT	CARBOHYD	419	419	POTENTIAL.
SQ	SEQUENCE	606 AA;	68126 MW; C07CCA06	CRC32;
Query Match				
Best Local Similarity				
Matches				
Db	456	IPAAPPTAASSTEEPAD-CDSYCKASKGKUKINMKYCKDYAVQIHILKAENADW-W	513	
Qy	153	IPLASSDHLPPATEAPKVEA-CCKNKDDDDMETLCKNDFALKIKV-REITYINRDT	210	
Db	514	KFTVNIIS--VYK-QG-SNR-LRGDOTLWVHANDIACKCPKVPKMK-YLLLGSTEDS	566	
Qy	211	KIILETKSKTIYKLVGSRDLKKS--VLWUKD-SLQCTCEEMDINAPYLVGMQKGQ	266	
RESULT				
ID	Y018.CABEL	STANDARD;	PRT;	465 AA.
AC	Q09280;			
DT	01-NOV-1995 (REL. 32, CREATED)			
DT	01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)			
DT	01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)			
DE	HYPOTHETICAL 53.8 KD PROTEIN C45G9.8 IN CHROMOSOME III.			
GN	C45G9.8			
OS	CAENORHABDITIS ELEGANS.			
OC	EUKARYOTA; METAZOA; ACCELOMATES; NEMATODA; SECERNENTEA; RHABDITIDA.			
OC	EUTHERIA; RODENTIA.			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=BRISTOL N2.			
RC	BENTLEY D., WATERSTON R.;			
RL	SUBMITTED (FEB-1995) TO EMBL/GENBANK/DBJ DATA BANKS.			
DR	EMBL: U21323; G687873; -			
DR	WORMPEP; C45G9.8; CE01859.			
KW	HYPOTHETICAL PROTEIN.			
FT	DOMAIN	139	145	POLY-THR.
SQ	SEQUENCE	465 AA;	53756 MW; 37C8121B	CRC32;
Query Match				
Best Local Similarity				
Matches				
Db	202	QLPRIEDONI-EV-TSRDIDEAD-MERICK-FAKEKINPITMAEPIDETKILLDRV	257	
Qy	160	HLPPATEAPKVEACNKKNDNDIMETLCKNDFAL-KIKVKEITY-INRDTKIILETK	217	
Db	258	TKKPYPLKYNKGL	272	
Qy	218	SKTIYKLVGVSERDL	232	
RESULT				
ID	CPD4.RAT	STANDARD;	PRT;	500 AA.
AC	R13108;			
DT	01-JAN-1990 (REL. 13, CREATED)			
DT	01-MAR-1992 (REL. 21, LAST SEQUENCE UPDATE)			
DT	01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)			
DE	CYTCHROME P450 IID4 (EC 1.14.14.1) (P450-DB4) (P450-CMF3)			
DE	(DEBRISOQUINE 4-HYDROXYLASE).			
GN	CYP2D14 OR CYP2D-4.			
OS	RATTUS NORVEGICUS (RAT).			
OC	EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;			
OC	EUTHERIA; RODENTIA.			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=SPRAGUE-DAWLEY; TISSUE=LIVER;			
RC	MEDLINE; 90189185.			
RX	MATSUNAGA E., UMEMO M., GONZALEZ F.J.;			
RL	J. MOL. EVOL. 30:155-169(1990).			

```
OS MUS MUSCULUS (MOUSE) .
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHYRIA; RODENTIA.
[1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 89384612.
RA ALCORTA D.A.; CREWS C.M., SWEET L.J., BANKSTON L., JONES S.W.,
RA ERIKSON R.L.;
RL COLL. CELL. BIOL. 9:3850-3859(1989).
CC -!- FUNCTION: PHOSPHORYLATES A WIDE RANGE OF SUBSTRATES INCLUDING
      RIBOSOMAL PROTEIN S6.
CC CC TISSUE SPECIFICITY: INTESTINE, THYMUS, AND LUNG.
CC -!- SIMILARITY: THE N-TERMINUS SHOWS HIGH SEQUENCE SIMILARITY TO THE
      REGIONS OF PROTEIN KINASE C, THE CATALYTIC SUBUNIT OF CAMP-
      DEPENDENT PROTEIN KINASE, AND GMP-DEPENDENT PROTEIN KINASE. THE
      REMINDER OF S6 KINASE II SHOWS HIGH SEQUENCE SIMILARITY TO THE
      CATALYTIC SUBUNIT OF PHOSPHORYLASE B KINASE.
CC EMBL; M28489; G556322; -.
DR DR PIR; B32571; B32571.
DR HSSP; P05132; ICTP.
DR MGd; MG1; I04558; RPS6KAL.
DR DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 2.
DR DR PROSITE; PS00108; PROTEIN_KINASE_ST; 2.
DR DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 2.
KW KW TRANSFERAS: SERINE/TREONINE-PROTEIN KINASE; ATP-BINDING;
    REPEAT; MULTIGENE FAMILY.
FT FT DOMAIN 52 310 PROTEIN KINASE 1.
FT FT DOMAIN 407 664 PROTEIN KINASE 2.
FT NP_BIND 68 76 ATP (BY SIMILARITY).
FT FT_BINDING 94 94 ATP (BY SIMILARITY).
FT ACT_SITE 187 187 BY SIMILARITY.
FT NP_BIND 413 421 ATP (BY SIMILARITY).
FT BINDING 436 436 ATP (BY SIMILARITY).
FT ACT_SITE 524 524 BY SIMILARITY.
SQ SEQUENCE 724 AA; 81594 MW; F961A261 CRC32;

Query Match          4.4%; Score 99; DB 1; Length 724;
Best Local Similarity 35.2%; Pred.No. 5,67e-02;
Matches 19; Conservative 13; Mismatches 18; Indels 4; Gaps

Db 489 KI-LRQKFSEAREASFVLTISKVEYLHSGQVVHRDLKPSNTLYVDGSGNPEC 541
   || ::: : ::::: ||| | | | | | | | | | | | | | | | | | | |
OY 197 KIKVKIEITYYNDTKIILETTSKTII-Y-KLVGVSRDLKK-S-VWLKLDSLOCTC 247

RESULT           9
ID VG07_BPT4 STANDARD; PRT; 1032 AA.
AC P19061;
DT 01-NOV-1990 (REL. 15, CREATED)
DD 01-NOV-1990 (REL. 16, LAST SEQUENCE UPDATE)
DE 01-NOV-1990 (REL. 16, LAST ANNOTATION UPDATE)
DE BASEPLATE STRUCTURAL PROTEIN GP7.
GN 7.
OS BACTERIOPHAGE T4.
OC VIRIDAE; DS-DNA NONENVELOPED VIRUSES; MYOVIRIDAE.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=D;
RX MEDLINE; 90384864.
RA EFIMOV V.P., PRILIPOV A.G., MESVANZHINOV V.V.;
RL NUCLEIC ACIDS RES 18:5313-5313(1990).
CC -!- FUNCTION: STRUCTURAL COMPONENT OF THE BASEPLATE.
DR EMBL; X15907; G15323; -.
DR PIR; JQ0657; G7BPT4.
KW STRUCTURAL PROTEIN.
SQ SEQUENCE 1032 AA; 119214 MW; 39F5B2D1 CRC32;

Query Match          4.4%; Score 98; DB 1; Length 1032;
Best Local Similarity 39.0%; Pred.No. 8,13e-02;
Matches 16; Conservative 8; Mismatches 15; Indels 2; Gaps

Db 368 MBSICKVFALLIGEVTFLANPNRKTSKII-DSADKGIVVLN 407
   ||:::| | | | | | | | | | | | | | | | | | | | | | | | | | | |
```

QY 186 METLCKNDNFALKI-KVKEITYINRDTKILETKSKTIYKLN 225

RESULT 10
ID ION3 CARAU STANDARD: PRT: 520 AA.
AC P18520;
DT 01-NOV-1990 (REL. 16, CREATED)
DT 01-NOV-1990 (REL. 16, LAST SEQUENCE UPDATE)
DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE INTERMEDIATE FILAMENT PROTEIN ON3.
OS CARASSIUS AURATUS (GOLDFISH).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; PISCES; GNATHOSTOMATA;
OC OSTEICHTHYES; ACTINOPTERYGII; CYPRINIFORMES.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-GLIAL CELL;
RX MEDLINE: 90180470.
RA GIORANO S., GLASGOW E., TESSER P., SCHECHTER N.;
RL NEURON 2:1507-1516(1989).
CC -!- FUNCTION: ONE OF THE NONNEURONAL PREDOMINANT INTERMEDIATE
FILAMENT PROTEINS OF THE VISUAL PATHWAY.
CC -!- SIMILARITY: SIGNIFICANT HOMOLOGY WITH MAMMALIAN TYPE II KERATIN
K8.
DR EMBL; M87773; -; NOT_ANNOTATED_CDS.
DR PIR; JS0291; JS0291.
DR PROSITE; PS00226; IF: 1.
KW INTERMEDIATE FILAMENT; COILED COIL; HEPTAD REPEAT PATTERN; VISION.
FT DOMAIN 1 108 HEAD.
FT DOMAIN 109 420 ROD.
FT DOMAIN 421 520 TAIL.
FT DOMAIN 109 144 COIL 1A.
FT DOMAIN 145 157 LINKER 1.
FT DOMAIN 158 253 COIL 1B.
FT DOMAIN 254 273 LINKER 12.
FT DOMAIN 274 420 COIL 2.
SQ SEQUENCE 520 AA; 57789 MW; 389CC03 CRC32;

Query Match 4.3%; Score 96; DB 1; Length 520;
Best Local Similarity 32.2%; Pred. No. 1.66e-01;
Matches 19; Conservative 14; Mismatches 24; Indels 2; Gaps 2;

DB 194 KNKYEDEIN-KRTEGENFVLTKKDVEAYMKVLEAKLSLSDSEINFLRIFEEIR 251
||| : : : | | | | : : : | : : : | : : : | : : : | : : :
QY 176 KNKNDNDNDIMETLCKNDNFALKIKVKEITYINR-DTKILETKSKTIYKLVGVSERDLK 233

RESULT 11
ID UD13-RAT STANDARD: PRT: 531 AA.
AC Q64637;
DT 01-NOV-1997 (REL. 35, CREATED)
DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE UDP-GLUCURONOSYLTRANSFERASE 1-3 PRECURSOR, MICROSMAL (EC 2.4.1.17)
DE (UDPGT) (UGT1+3) (UGT1-03) (UGT1A3) (B3).
GN UGT1.
OS RATTUS NORVEGICUS (RAT).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; RODENTIA.
RN [1]
RP SEQUENCE OF 1-286 FROM N.A.
RC STRAIN-WISTAR;
RX MEDLINE: 95332265.
RA EMI Y., IKUSHIRO S.I., IYANAGI T.;
RL J. BIOCHEM. 117:392-399(1995).
RN [2]
RP SEQUENCE OF 287-531 FROM N.A.
RC TISSUE-LIVER;
RX MEDLINE: 90274676.
RA SATO H., KOINAI O., TANABE K., KASHIWAMATA S.;
RL BIOCHEM. BIOPHYS. RES. COMMUN. 169:260-264(1990).
CC -!- FUNCTION: UDPGT IS OF MAJOR IMPORTANCE IN THE CONJUGATION AND
SUBSEQUENT ELIMINATION OF POTENTIALLY TOXIC XENOBIOTICS AND
ENDOGENOUS COMPOUNDS.

CC -!- CATALYTIC ACTIVITY: UDP-GLUCURONATE + ACCEPTOR -> UDP + ACCEPTOR
CC BETA-D-GLUCURONOSIDE.
CC -!- SUBCELLULAR LOCATION: THE UGT1 GENE ENCODES FOR MANY DIFFERENT
CC -!- ALTERNATIVE PRODUCTS: THE UGT1 GENE ENCODES FOR MANY DIFFERENT
CC UDPGT ISOZYMES WHICH HAVE A DIFFERENT N-TERMINAL DOMAIN AND A
CC COMMON C-TERMINAL DOMAIN OF 245 RESIDUES.
CC -!- SIMILARITY: BELONGS TO THE UDP-GLYCOSYLTRANSFERASE FAMILY.
DR EMBL; D38067; G940422; -.
DR EMBL; M34007; G207579; ALT TERM.
KW PROSITE; PS00375; UDPGT: 1.
KW TRANSFERASE: GLYCOSYLTRANSFERASE; GLYCOPROTEIN; TRANSMEMBRANE; SIGNAL;
KW MULTIGENE FAMILY; MICROSMAL; ALTERNATIVE SPLICING.
FT SIGNAL 1 25 POTENTIAL.
FT CHAIN 26 531 UDP-GLUCURONOSYLTRANSFERASE 1-3.
FT TRANSMEM 489 505 POTENTIAL.
FT CARBOHYD 116 116 POTENTIAL.
FT CARBOHYD 139 139 POTENTIAL.
FT CARBOHYD 293 293 POTENTIAL.
FT CARBOHYD 431 431 POTENTIAL.
SQ SEQUENCE 531 AA; 60140 MW; C3B2A846 CRC32;

Query Match 4.2%; Score 94; DB 1; Length 531;
Best Local Similarity 35.7%; Pred. No. 3.35e-01;
Matches 15; Conservative 15; Mismatches 7; Indels 5; Gaps 5;

DB 214 PVP-WMYLCH-VNYGSLARLASDLQREVSVVEILRHASVWL 253

QY 42 PIPANLQCHGIEYQNM-RLP-NLLGHE-TMKVEVLEQAGAMI 80

RESULT 12
ID YMB1-YEAST STANDARD: PRT: 1562 AA.
AC Q04781; Q04029;
DT 01-NOV-1997 (REL. 35, CREATED)
DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE HYPOTHETICAL 180.2 KD PROTEIN IN FAA4-HOR7 INTERGENIC REGION.
GN YMB247C OR YMB408.09C OR YMB920.01C.
OS SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).
OC EUKARYOTA; FUNGI; ASCOMYCOTINA; HEMIASCOMYCETES.
RN [1]
RP SEQUENCE OF 1-956 FROM N.A.
RC STRAIN-S288C / AB972;
RA HUNT S., BOWMAN S., BARRELL B.G., RAJANDREAM M.A., WALSH S.V.;
RL SUBMITTED (MAR-1995) TO EMBL/GENBANK/DBJ DATA BANKS.
RN [2]
RP SEQUENCE OF 950-1562 FROM N.A.
RC STRAIN-S288C / AB972;
RA GENTLES S., BOWMAN S., BARRELL B.G., RAJANDREAM M.A., WALSH S.V.;
RL SUBMITTED (MAR-1995) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL; 248639; G732925; -.
DR EMBL; 248756; G736313; -.
KW HYPOTHETICAL PROTEIN.
SQ SEQUENCE 1562 AA; 180185 MW; CEE45672 CRC32;

Query Match 4.1%; Score 92; DB 1; Length 1562;
Best Local Similarity 38.8%; Pred. No. 6.68e-01;
Matches 19; Conservative 9; Mismatches 21; Indels 0; Gaps 0;

DB 1397 KNEFDINDSKMDRLTSNDALTIKLNITNEVKASYLDDOKLEISFKL 1445
||| : : : | | | | : : : | : : : | : : : | : : : | : : :
QY 176 KNKNDNDNDIMETLCKNDNFALKIKVKEITYINRDTKILETKSKTIYKL 224

RESULT 13
ID GYRB-ACIS3 STANDARD: PRT: 216 AA.
AC Q44273; Q60168;
DT 01-NOV-1997 (REL. 35, CREATED)
DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE DNA GYRASE SUBUNIT B (EC 5.99.1.3) (FRAGMENTS).
GN GYRB.
OS ACINETOBACTER SP. (STRAIN ATCC 17924 / CIP 70.12).

```

CC PROKARYOTA; GRACILICUTES; SCOTOBACTERIA; AEROBIC RODS AND COCCI;
CC NEISSERIACEAE.
CC [1]
CC SEQUENCE FROM N.A.
CC MEDLINE: 9708999.
CC YAMAMOTO S., HARAYAMA S.;
CC INT. J. SYST. BACTERIOL. 46:506-511(1996).
CC -!- FUNCTION: DNA GYRASE NEGATIVELY SUPERCOILS CLOSED CIRCULAR DOUBLE-
CC STRANDED DNA IN AN ATP-DEPENDANT MANNER AND ALSO CATALYZES THE
CC INTERCONVERSION OF OTHER TOPOLOGICAL ISOMERS OF DOUBLE-STRANDED
CC DNA RINGS, INCLUDING CATENANES AND KNOTTED RINGS.
CC -!- CATALYTIC ACTIVITY: ATP-DEPENDENT BREAKAGE, PASSAGE AND REJOINING
CC OF DOUBLE-STRANDED DNA.
CC -!- SUBUNIT: MADE UP OF TWO CHAINS. THE A CHAIN IS RESPONSIBLE FOR DNA
CC BREAKAGE AND REJOINING; THE B CHAIN CATALYZES ATP HYDROLYSIS. THE
CC ENZYME FORMS AN A2B2 TETRAMER.
CC EMBL: D73436; G1322131; -.
CC PROSITE: D73421; G1322101; -.
CC TOPOISOMERASE; ISOMERASE; ATP-BINDING; ANTIBIOTIC RESISTANCE.
CC NON_TER 1
CC NON_CONS 116 117
CC NON_TER 216 216
CC SEQUENCE 216 AA; 23775 MW; E572EE0 CRC32;

Query Match 4.0%; Score 90; DB 1; Length 216;
Best Local Similarity 26.3%; Pred. No. 1.31e+00;
Matches 20; Conservative 23; Mismatches 26; Indels 7; Gaps 7;

Db 79 LARRRLELFLNAGVRIVLRERVALEHFDLEVLGSEKSL-DIAGLPGLK-A-DCEKD 136
Qy 196 LKIKVKEITYINRDKILL-ETK-S-KTIYKLN-GVSEKDLKSLVLMKDSLOCTCEMN 251
Db 137 PALSELXLVEGDSAGG 152
Qy 252 -DINAPYLVMQKQGG 266

RESULT 14
ID VIRF_SHIDY STANDARD; PRT; 262 AA.
AC Q04248:
DT 01-JUN-1994 (REL. 29, CREATED)
DT 01-JUN-1994 (REL. 29, LAST SEQUENCE UPDATE)
DT 01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)
DE VIRULENCE REGULON TRANSCRIPTIONAL ACTIVATOR VIRF.
GN VIRF.
OS SHIGELLA DYSENTERIAE, SHIGELLA FLEXNERI, AND SHIGELLA SONNEI.
OG PLASMID 210 KB INVASION.
CC PROKARYOTA; GRACILICUTES; SCOTOBACTERIA; FACULTATIVELY ANAEROBIC RODS;
CC ENTEROBACTERIACEAE.
CC [1]
CC SEQUENCE FROM N.A.
CC SPECIES=S.DYSENTERIAE; STRAIN=CG097;
CC YAO R., REDDY L.V., PALCHANDHURI S.;
CC SUBMITTED (MAR-1991) TO EMBL/GENBANK/DBJ DATA BANKS.
CC [2]
CC SEQUENCE FROM N.A.
CC SPECIES=S.FLEXNERI; STRAIN=2A;
CC MEDLINE: 87032409.
CC SAKAI T., SASAKAWA C., MAKINO S., YOSHIKAWA M.;
CC INFECT. IMMUN. 54:395-402(1986).
CC [3]
CC SEQUENCE FROM N.A.
CC SPECIES=S.SONNEI;
CC MEDLINE: 89212861.
CC KATO J.I., ITO K.I., NAKAMURA A., WATANABE H.;
CC INFECT. IMMUN. 57:1391-1398(1989).
CC [4]
CC SIMILARITY TO ARAC FAMILY.
CC MEDLINE: 92326642..
CC DORMAN C.J.;
CC MOL. MICROBIOL. 6:1575-1575(1992).
CC -!- FUNCTION: TRANSCRIPTIONAL ACTIVATOR OF THE VIRB GENE WHICH IS

```

```

CC ITSELF AN ACTIVATOR OF THE IPABCD VIRULENCE REGULON.
CC -!- SIMILARITY: BELONGS TO THE ARAC/XVLS FAMILY OF TRANSCRIPTIONAL
CC REGULATORS.
CC EMBL: X58464; G46954; -.
CC MEDLINE: M29172; G152801; -.
CC EMBL: X16661; G47067; -.
CC DR PIR: S14646; S14646.
CC DR PIR: A47605; A47605.
CC DR PIR: A60105; A60105.
CC PROSITE: PS00041; HTH_ARAC_FAMILY_1; 1.
CC PROSITE: PS01124; HTH_ARAC_FAMILY_2; 1.
CC KW PLASMID; VIRULENCE; TRANSCRIPTION REGULATION; ACTIVATOR; DNA-BINDING.
CC FT DNA_BIND 177 196 H-T-H MOTIF (BY SIMILARITY).
CC SEQUENCE 262 AA; 30549 MW; 11339B38 CRC32;

Query Match 4.0%; Score 89; DB 1; Length 262;
Best Local Similarity 25.5%; Pred. No. 1.83e+00;
Matches 13; Conservative 21; Mismatches 15; Indels 2; Gaps 2;

Db 37 TLAIDEGQIAFIERNQINVSIIKSDSINPFETISIDRNLLLSIIIMPEPI 87
Qy 195 ALKIKVKEITYINRDKILL-ETK-S-KTIYKLVGS-ERDLKKSVLWKDSL 243

RESULT 15
ID KSGA_MYCPN STANDARD; PRT; 263 AA.
AC P75113:
DT 01-NOV-1997 (REL. 35, CREATED)
DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE DIMETHYLADENOSINE TRANSFERASE (EC 2.1.1.-) (S-ADENOSYLMETHIONINE-6-N',
DE N'-ADENOSYL(RRNA) DIMETHYLTRANSFERASE) (16S RNA DIMETHYLASE) (HIGH
DE LEVEL KASUGAMYCIN RESISTANCE PROTEIN KSGA) (KASUGAMYCIN
DE DIMETHYLTRANSFERASE).
GN KSGA.
OS MYCOPLASMA PNEUMONIAE.
CC PROKARYOTA; TENERICUTES; MOLLICUTES; MYCOPLASMA; MYCOPLASMATALES;
CC MYCOPLASMATACEAE.
CC [1]
CC SEQUENCE FROM N.A.
CC STRAIN=ATCC 29342 / M129;
CC MEDLINE: 97105885.
CC RA HIMMELREICH R., HILBERT H., PLAGENS H., PIRKL E., LI B.-C.,
CC HERRMANN R.;
CC NUCLEIC ACIDS RES. 24:4420-4449(1996).
CC -!- FUNCTION: SPECIFICALLY DIMETHYLATES TWO ADJACENT ADENOSINES IN THE
CC LOOP OF A CONSERVED HAIRPIN NEAR THE 3'-END OF 16S RNA IN THE 30S
CC PARTICLE. ITS INACTIVATION LEADS TO KASUGAMYCIN RESISTANCE
CC (BY SIMILARITY).
CC -!- SIMILARITY: STRONG, TO OTHER BACTERIAL KSGA.
CC -!- SIMILARITY: BELONGS TO THE RNA ADENINE N-6-METHYLTRANSFERASE
CC FAMILY.
CC EMBL: AE000017; G1673824; -.
CC PROSITE: PS01131; RNA_A_DIMETH; 1.
CC KW MRNA PROCESSING; TRANSFERASE; METHYLTRANSFERASE;
CC ANTIBIOTIC RESISTANCE.
CC SEQUENCE 263 AA; 29794 MW; 357AA00D CRC32;

Query Match 4.0%; Score 90; DB 1; Length 263;
Best Local Similarity 37.0%; Pred. No. 1.31e+00;
Matches 20; Conservative 11; Mismatches 20; Indels 3; Gaps 3;

Db 52 KLQLPYHGIEL-DKRLAEVLLVNEILTEQLTGIDALKONLDQYFDPDTPLLCG 104
Qy 46 NLQL-CHGIEYQNMRLPN-LLGHETMKVELEGAGAWIPLVMKQCHPDTKKFLCS 97

Search completed: Tue Aug 4 10:34:10 1998
Job time : 26 secs.

```

M P S R E L H (TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPsrch_pp protein - protein database search, using Smith-Waterman algorithm
Run on: Tue Aug 4 10:34:28 1998; MasPar time 19.77 Seconds
Tabular output not generated. 628.445 Million cell updates/sec

Title: >US-08-848-439-2
Description: (1-295) from US08848439.pep
Perfect Score: 2234
Sequence: 1 MLOGPGSLLLLFLASHCCLG.....WKGOREFKRSIRKLCQ 295

Scoring table: PAM 150
Gap 11

Searched: 140555 seqs, 42109429 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: sptribl6
1:sp_fungi 2:sp_human 3:sp_invertebrate 4:sp_mammal
5:sp_mhc 6:sp_organelle 7:sp_phage 8:sp_plant
9:sp_bacteria 10:sp_rodent 11:sp_virus 12:sp_invertebrate
13:sp_unclassified

Statistics: Mean 44.964; Variance 71.174; scale 0.632
Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES					Pred. No.	
Result No.	Score	Query Match	Length	ID	Description	
1	2190	98.0	295	10	008862	SECRETED FRIZZLED REL 0.00e+00
2	2188	97.9	295	10	035297	SECRETED APOPTOSIS REL 0.00e+00
3	2185	97.8	295	10	P97299	STROMAL CELL DERIVED F 0.00e+00
4	1538	68.8	206	2	014778	SECRETED APOPTOSIS REL 0.00e+00
5	913	40.9	314	2	014779	SECRETED APOPTOSIS REL 5.25e-199
6	911	40.8	313	2	000546	SECRETED FRIZZLED-RELA 1.75e-198
7	902	40.4	308	4	019116	FRZA PRECURSOR. 3.93e-196
8	901	40.3	317	2	014780	SECRETED APOPTOSIS REL 7.17e-196
9	890	39.8	314	10	008861	SECRETED FRIZZLED REL 5.34e-193
10	729	32.6	307	12	042397	CRESCENT. 3.39e-151
11	400	17.9	588	3	016147	MON-5. 4.73e-68
12	396	17.7	572	10	061090	FRIZZLED HOMOLOG 7 (TR 4.50e-67
13	394	17.6	591	2	000144	FRIZZLED HOMOLOG. 1.39e-66
14	394	17.6	685	10	061091	FRIZZLED HOMOLOG 8 (TR 1.39e-66
15	386	17.3	565	2	014332	FRIZZLED GENE PRODUCT. 1.24e-64
16	386	17.3	570	10	008464	FRIZZLED PROTEIN HOMOL 1.24e-64
17	386	17.3	641	10	008463	FRIZZLED PROTEIN HOMOL 1.24e-64
18	374	16.7	585	2	013467	TRANSMEMBRANE RECEPTOR 1.03e-61
19	369	16.5	537	10	061088	FRIZZLED HOMOLOG 4 (TR 1.68e-60
20	357	16.0	525	3	010662	FRIZZLED PROTEIN HOMOL 1.33e-57

21	355	15.9	694	3	Q94916	DFZ2.	4.05e-57
22	347	15.5	586	3	Q24760	TISSUE POLARITY PROTEI	3.41e-55
23	337	15.1	558	3	Q94132	TRANSMEMBRANE RECEPTOR	8.51e-53
24	317	14.2	666	10	Q61086	FRIZZLED-3.	4.93e-48
25	309	13.8	664	12	O42579	FRIZZLED 3. PROTEIN.	3.85e-45
26	295	13.2	605	3	O16190	FZ2E11.3 PROTEIN.	7.55e-43
27	293	13.1	325	4	O95117	FRZB PRECURSOR.	2.22e-42
28	291	13.0	325	2	O00181	FRITZ.	6.51e-42
29	289	12.9	325	2	O95686	FRZB PRECURSOR.	1.91e-41
30	289	12.9	325	2	O92765	FRZB PRECURSOR.	1.91e-41
31	284	12.7	709	10	O61089	FRIZZLED HOMOLOG 6 (TR	2.80e-40
32	281	12.6	348	10	O35222	FRIZZLED RELATED PROTE	1.39e-39
33	275	12.3	323	10	P97401	SECRETED FRIZZLED-RELA	3.43e-38
34	274	12.3	368	2	O14877	FRPHE.	5.84e-38
35	267	12.0	318	12	P79936	FRZB PRECURSOR.	2.40e-36
36	267	12.0	319	12	P79993	FRZB PRECURSOR.	2.40e-36
37	240	10.7	261	10	O08570	FRITZ (FRAGMENT).	3.38e-30
38	195	8.7	641	2	O00520	CARBOXYPEPTIDASE Z PRE	2.67e-20
39	178	8.0	505	10	O60569	PROCOLLAGEN, TYPE XVII	1.04e-16
40	178	8.0	562	10	O60572	PROCOLLAGEN, TYPE XVII	1.04e-16
41	178	8.0	1774	10	Q62001	ALPHA-1(XVII) COLLAGE	1.04e-16
42	125	5.6	602	12	O42203	NETRIN 1.	2.58e-06
43	124	5.6	794	12	O42224	SMOOTHENED (FRAGMENT).	3.91e-06
44	122	5.5	787	2	O98635	SMOOTHENED.	8.94e-06
45	122	5.5	793	10	P97698	SMOOTHENED.	8.94e-06

ALIGNMENTS

RESULT 1
ID 008862; PRELIMINARY; PRT: 295 AA.
AC 008862;
DT 01-JUL-1997 (TREMBLREL. 04, CREATED)
DT 01-JUL-1997 (TREMBLREL. 04, LAST SEQUENCE UPDATE)
DT 01-JUL-1997 (TREMBLREL. 04, LAST ANNOTATION UPDATE)
DE SECRETED FRIZZLED RELATED PROTEIN SFRP-2.
GN SFRP2.
OS MUS MUSCULUS (MOUSE).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; RODENTIA.
RN [1]
RP SEQUENCE FROM N.A.
RA RATTNER A., HSIEH J.C., SMALLWOOD P.M., GILBERT D.J.,
RA COPELAND N.G., JENKINS N.A., NATHANS J.;
RL PROC. NATL. ACAD. SCI. U.S.A. 94:2859-2863(1997).
DR EMBL: U88567; G1946343; -
SQ SEQUENCE 295 AA; 33469 MW; 23680DDA CRC32;

Query Match					98.0%; Score 2190; DB 10; Length 295;
Best Local Similarity					98.0%; Pred. No. 0.00e+00;
Matches 289; Conservative 3; Mismatches 3; Indels 0; Gaps 0;					
Db	1	MPRGASILLILLVASHCCGLGSARGFLFGQDPFSYKRSNCKP	IPANLQICHGIEYQNMRL	60	
Qy	1	MLQPGSLLLLFLASHCCGLGSARGFLFGQDPFSYKRSNCKP	IPANLQICHGIEYQNMRL	60	
Db	61	PNLLGHETMKEVLEQAGAWIPVMKQCHDPTKKFLCSLFAPVCLDDDET	IQPCHSLCQV	120	
Qy	61	PNLLGHETMKEVLEQAGAWIPVMKQCHDPTKKFLCSLFAPVCLDDDET	IQPCHSLCQV	120	
Db	121	VKDRCAPVMSAFGFPWPMLECDRFPQDNDLCIPASSDHLLPATEAPKVC	EACKTKNE	180	
Qy	121	VKDRCAPVMSAFGFPWPMLECDRFPQDNDLCIPASSDHLLPATEAPKVC	EACKTKNE	180	
Db	181	DNDIMETLCKNDNFALKIKVKETIYNRDKIILETKSTIYKLVGVSERDLK	KSVLWLK	240	
Qy	181	DNDIMETLCKNDNFALKIKVKETIYNRDKIILETKSTIYKLVGVSERDLK	KSVLWLK	240	
Db	241	DSLQCTCEEMNDINAPYLVMGQKGELVITSVKRWQKQGRFKRSIRKLCQ	295		
Qy	241	DSLQCTCEEMNDINAPYLVMGQKGELVITSVKRWQKQGRFKRSIRKLCQ	295		

```

RESULT 2
ID O35297 PRELIMINARY; PRT: 295 AA.
AC O35297;
DT 01-JAN-1998 (TREMBLREL. 05, CREATED)
DT 01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)
DT 01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)
DE SECRETED APOPTOSIS RELATED PROTEIN 1.
GN SARPI.
OS MUS MUSCULUS (MOUSE).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; RODENTIA.
RN [1]
RP SEQUENCE FROM N.A.
RA MELKONYAN H., CHANG W.C., SHAPIRO J.P., MAHADEVAPPA M., FITZPATRICK P.A.,
RA FITZPATRICK P.A., KIEFER M.C., TOMEI D.L., UMANSKY S.R.;
RL SUBMITTED (SEP-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL; AF017989; G2415421; -.
SQ SEQUENCE 295 AA; 33483 MW; 0E4C8A8E CRC32;

Query Match 97.9%; Score 2188; DB 10; Length 295;
Best Local Similarity 97.6%; Pred. No. 0.00e+00;
Matches 288; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Db 1 MPRGPASLLLVASHCCLSGARGFLFGQPDFSYKRNCKRPIPANLQCHGIEYQNMRL 60
QY 1 MLOQPGSLLLVASHCCLSGARGFLFGQPDFSYKRNCKRPIPANLQCHGIEYQNMRL 60
Db 61 PNLLGHETMKEVLEQAGAWIPLVMKQCHPDTKKFLCSLFAFVCLDDLDDETQPCHSLCQ 120
QY 61 PNLLGHETMKEVLEQAGAWIPLVMKQCHPDTKKFLCSLFAFVCLDDLDDETQPCHSLCQ 120
Db 121 VKDRCAPVMSAFGFPWDMLECDRFPQDNDLCIPLASSDHLHPATEAPKVCACKTKNE 180
QY 121 VKDRCAPVMSAFGFPWDMLECDRFPQDNDLCIPLASSDHLHPATEAPKVCACKTKNE 180
Db 181 DDNDIMETLCKNDPALKIKVKEITYINRDKIILETSKTIYKLVGVSERDLKKSVLWLK 240
QY 181 DDNDIMETLCKNDPALKIKVKEITYINRDKIILETSKTIYKLVGVSERDLKKSVLWLK 240
Db 241 DSLOQTCCEMDINAPYLVMGQKGGELVITSVKRWQKGOREFKRISIRKLQ 295
QY 241 DSLOQTCCEMDINAPYLVMGQKGGELVITSVKRWQKGOREFKRISIRKLQ 295

Query Match 97.8%; Score 2185; DB 10; Length 295;
Best Local Similarity 97.6%; Pred. No. 0.00e+00;
Matches 288; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Db 1 MPRGPASLLLVASHCCLSGARGFLFGQPDFSYKRNCKRPIPANLQCHGIEYQNMRL 60
QY 1 MLOQPGSLLLVASHCCLSGARGFLFGQPDFSYKRNCKRPIPANLQCHGIEYQNMRL 60
Db 61 PNLLGHETMKEVLEQAGAWIPLVMKQCHPDTKKFLCSLFAFVCLDDLDDETQPCHSLCQ 120
QY 61 PNLLGHETMKEVLEQAGAWIPLVMKQCHPDTKKFLCSLFAFVCLDDLDDETQPCHSLCQ 120
Db 121 VKDRCAPVMSAFGFPWDMLECDRFPQDNDLCIPLASSDHLHPATEAPKVCACKTKNE 180
QY 121 VKDRCAPVMSAFGFPWDMLECDRFPQDNDLCIPLASSDHLHPATEAPKVCACKTKNE 180
Db 181 DDNDIMETLCKNDPALKIKVKEITYINRDKIILETSKTIYKLVGVSERDLKKSVLWLK 240
QY 181 DDNDIMETLCKNDPALKIKVKEITYINRDKIILETSKTIYKLVGVSERDLKKSVLWLK 240
Db 241 DSLOQTCCEMDINAPYLVMGQKGGELVITSVKRWQKGOREFKRISIRKLQ 295
QY 241 DSLOQTCCEMDINAPYLVMGQKGGELVITSVKRWQKGOREFKRISIRKLQ 295

RESULT 3
ID P97299 PRELIMINARY; PRT: 295 AA.
AC P97299;
DT 01-MAY-1997 (TREMBLREL. 03, CREATED)
DT 01-MAY-1997 (TREMBLREL. 03, LAST SEQUENCE UPDATE)
DT 01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)
DE STROMAL CELL DERIVED FACTOR 5 (SDF5).
GN SDF5.
OS MUS MUSCULUS (MOUSE).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; RODENTIA.
RN [1]
RP SEQUENCE FROM N.A.
RA SHIROZU M., TADA H., TASHIRO K., NAKAMURA T., LOPEZ N.D.,
RA NAZAREM M., HAMADA T., SATO T., NAKANO T., HONJO T.;
RL GENOMICS 37:273-280(1996).
DR EMBL; D50462; G1747302; -.
DR MGD; MGI:108078; SDF5.
SQ SEQUENCE 295 AA; 33501 MW; 329888C8 CRC32;

Query Match 97.8%; Score 2185; DB 10; Length 295;
Best Local Similarity 97.6%; Pred. No. 0.00e+00;
Matches 288; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Db 1 MPRGPASLLLVASHCCLSGARGFLFGQPDFSYKRNCKRPIPANLQCHGIEYQNMRL 60
QY 1 MLOQPGSLLLVASHCCLSGARGFLFGQPDFSYKRNCKRPIPANLQCHGIEYQNMRL 60
Db 61 PNLLGHETMKEVLEQAGAWIPLVMKQCHPDTKKFLCSLFAFVCLDDLDDETQPCHSLCQ 120

```

```

QY 61 PNLLGHETMKEVLEQAGAWIPLVMKQCHPDTKKFLCSLFAFVCLDDLDDETQPCHSLCQ 120
Db 121 VKDRCAPVMSAFGFPWDMLECDRFPQDNDLCIPLASSDHLHPATEAPKVCACKTKNE 180
QY 121 VKDRCAPVMSAFGFPWDMLECDRFPQDNDLCIPLASSDHLHPATEAPKVCACKTKNE 180
Db 181 DDNDIMETLCKNDPALKIKVKEITYINRDKIILETSKTIYKLVGVSERDLKKSVLWLK 240
QY 181 DDNDIMETLCKNDPALKIKVKEITYINRDKIILETSKTIYKLVGVSERDLKKSVLWLK 240
Db 241 DSLOQTCCEMDINAPYLVMGQKGGELVITSVKRWQKGOREFKRISIRKLQ 295
QY 241 DSLOQTCCEMDINAPYLVMGQKGGELVITSVKRWQKGOREFKRISIRKLQ 295

RESULT 4
ID O14778 PRELIMINARY; PRT: 206 AA.
AC O14778;
DT 01-JAN-1998 (TREMBLREL. 05, CREATED)
DT 01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)
DT 01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)
DE SECRETED APOPTOSIS RELATED PROTEIN 1 (FRAGMENT).
GN SARPI.
OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; PRIMATES.
RN [1]
RP SEQUENCE FROM N.A.
RA MELKONYAN H., CHANG W.C., SHAPIRO J.P., MAHADEVAPPA M., FITZPATRICK P.A.,
RA KIEFER M.C., TOMEI D.L., UMANSKY S.R.;
RL SUBMITTED (SEP-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL; AF017986; G2415415; -.
FT NON_TER 206
SQ SEQUENCE 206 AA; 23215 MW; F465B68B CRC32;

Query Match 68.8%; Score 1538; DB 2; Length 206;
Best Local Similarity 98.6%; Pred. No. 0.00e+00;
Matches 205; Conservative 0; Mismatches 1; Indels 2; Gaps 1;

Db 1 MLOQPGSLLLVASHCCLSGARGFLFGQPDFSYKRNCKRPIPANLQCHGIEYQNMRL 60
QY 1 MLOQPGSLLLVASHCCLSGARGFLFGQPDFSYKRNCKRPIPANLQCHGIEYQNMRL 60
Db 61 PNLLGHETMKEVLEQAGAWIPLVMKQCHPDTKKFLCSLFAFVCLDDLDDETQPCHSLCQ 120
QY 61 PNLLGHETMKEVLEQAGAWIPLVMKQCHPDTKKFLCSLFAFVCLDDLDDETQPCHSLCQ 120
Db 121 VKDRCAPVMSAFGFPWDMLECDRFPQDNDLCIPLASSDHLHPATEAPKVCACKTKNE 178
QY 121 VKDRCAPVMSAFGFPWDMLECDRFPQDNDLCIPLASSDHLHPATEAPKVCACKTKNE 180
Db 179 DDNDIMETLCKNDPALKIKVKEITYINR 206
QY 181 DDNDIMETLCKNDPALKIKVKEITYINR 208

RESULT 5
ID O14779 PRELIMINARY; PRT: 314 AA.
AC O14779;
DT 01-JAN-1998 (TREMBLREL. 05, CREATED)
DT 01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)
DT 01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)
DE SECRETED APOPTOSIS RELATED PROTEIN 2.
GN SARPI.
OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; PRIMATES.
RN [1]
RP SEQUENCE FROM N.A.
RA MELKONYAN H., PROCHAZKA V., CHANG W.C., UMANSKY S.R.;
RL SUBMITTED (SEP-1997) TO EMBL/GENBANK/DBJ DATA BANKS.

```

ID	Q19116	PRELIMINARY;	PRT;	308 AA.
AC	Q19116;			
DT	01-JAN-1998	(TREMBLREL. 05, CREATED)		
DT	01-JAN-1998	(TREMBLREL. 05, LAST SEQUENCE UPDATE)		
DT	01-JAN-1998	(TREMBLREL. 05, LAST ANNOTATION UPDATE)		
DE	FRZA, PRECURSOR.			
OS	BOS TAURUS (BOVINE).			
OC	EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;			
OC	EUTHERIA; ARTIODACTYLA.			
OC	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=HEART;			
RA	DUPLAA C., D'AMORE P.A.;			
RL	SUBMITTED (AUG-1997) TO EMBL/GENBANK/DBJ DATA BANKS.			
DR	EMBL; U85945; G2337937;			
KW	SIGNAL.			
FT	SIGNAL	5	22	POTENTIAL.
SEQ	SEQUENCE	308 AA;	34763 MR;	9B7AA92F CRC32;
	Query Match	40.4%;	Score 902;	DB 4;
	Best Local Similarity	42.2%;	Pred. No. 3.93e-196;	
	Matches	111;	Conservative	66; Mismatches 79; Indels 7; Gaps 5;
Db	45	FTYKPPQCDIPADRLCHLVNGYKRMVLPNLLHEHETMAEVKQQAQSSWVPLLNKNCHIGTQ	104	
Qy	33	FSYKRSNCKPIPANQLQCHGIEYQNMRPLNLLGHETMKEVLEQAGAWIPLVMKQCHPDTK	92	
Db	105	VFLCSLFPAYCLD--RPIYPCRWLCEAVRDSCEPVYMOFFGVYVPEMLKCDKFP-EGDYC	160	
Qy	93	KFLCSLFPAYCLDLDLDTIQCHSLCVQVRCAPVMSAFGFPWPNMLECDRFPQNDLJC	152	
Db	161	IAMTPPNATEASRQGTTCVPCPDNKLKSEA-IIEHLCASEFALRMKIKREVKNKGDKKI	219	
Qy	153	IPLASSDHLPLATEAPKVCCEACKNDDNDIMETLCKNDPALKIKVKEITYINRDTKI	212	
Db	220	V-PKKKKPL-KLGPIKKELKLVLYLKNAGDCPCQHLQDNLSHHFLINGRKVKYSQVLLTA	277	
Qy	213	ILETKSTIYKLVGVSERDLKKSVLWKLSQDTCCEMNDINAPYLVMGQKOGGELVITS	272	
Db	278	IHKWDKKNEKFEKTMKKMKNHCEC	300	
Qy	273	VKRWQKQGRFKRISRSIRKLQC	295	
RESULT	.8			
ID	Q14780	PRELIMINARY;	PRT;	317 AA.
AC	Q14780;			
DT	01-JAN-1998	(TREMBLREL. 05, CREATED)		
DT	01-JAN-1998	(TREMBLREL. 05, LAST SEQUENCE UPDATE)		
DT	01-JAN-1998	(TREMBLREL. 05, LAST ANNOTATION UPDATE)		
DE	SECRETED APOPTOSIS RELATED PROTEIN 3.			
GN	SARP3.			
OS	HOMO SAPIENS (HUMAN).			
OC	EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;			
OC	EUTHERIA; PRIMATES.			
OC	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=PANCREAS;			
RA	MELKONYAN H., CHANG W.C., SHAPIRO J.P., MAHADEVAPPA M., FITZPATRICK P.A.;			
RA	KIEFER M.C., TOMEI D.L., UMANSKY S.R.;			
RL	SUBMITTED (SEP-1997) TO EMBL/GENBANK/DBJ DATA BANKS.			
DR	EMBL; AF017988; G2415419;			
SEQ	SEQUENCE	317 AA;	35577 MR;	2CD69373 CRC32;
	Query Match	40.3%;	Score 901;	DB 2;
	Best Local Similarity	45.2%;	Pred. No. 7.17e-196;	
	Matches	114;	Conservative	58; Mismatches 74; Indels 6; Gaps 4;
Db	46	SYKPPQCDIPADRLPLCHTVNGYKRMVLPNLLHEHESLAEVKKQAQSSWVPLLNKCRHSDTO	105	
Qy	34	SY-KRSNCKPIPANQLQCHGIEYQNMRPLNLLGHETMKEVLEQAGAWIPLVMKQCHPDTK	92	
Db	106	VFLCSLFPAYCLD--RPIYPCRSLSCEAVRAGCAPLMEAYGFPWPEMLKCHKFPPLDNLJC	162	

Qy 162 LPA 164
Search completed: Tue Aug 4 10:35:55 1998
Job time : 87 secs.

THIS PAGE BLANK (USPTO)

QY 24 PANLQCHGIEYQNMRLPNLLGHETMKVEVLEQAGAWIPLVMKQCHPDTKKFLCSLFAFVC 83
 Db 111 -tvldqaiapprscetargcealmnkfgqwpelrcenfphvgageicvvgntsd 167
 QY 84 LDDLDLTIQCHSLCVQVKDRCAPVMSAFGFWPDMLECDREFP-ODN-DLICPIASSD 139

RESULT 2
 ID W31274 standard; Protein: 685 AA.
 AC W31274:
 DT 27-APR-1998 (first entry)
 DE Mouse frizzled-8 protein Mfz8 (Wnt receptor).
 KW Wnt receptor; mouse frizzled-8 protein; Mfz8 gene;
 KW signal transduction; cancer; cell growth; cell proliferation.
 OS Mus musculus.
 PN W09739357-A1.
 PD 23-OCT-1997.
 PF 11-APR-1997; U06049.
 PR 12-APR-1996; US-015307.
 PA (UYJO) UNIV LELAND STANFORD JUNIOR.
 PA (STRD) UNIV LELAND STANFORD JUNIOR.
 PI Andrew D, Bhanot P, Brink M, Hsieh J, Nathans J,
 PI Nusse R, Samos CH, Wangy;
 DR WPI: 97-526631/48.
 DR N-PSDB; T89892.

PT Identification of Wnt receptor binding modulators - useful for
 PT treatment of cancer and growth, development or proliferation related
 PT disorders
 PS Disclosure: Page 48-50; 61pp; English.
 CC This protein comprises the mouse transmembrane receptor,
 CC frizzled-8 (Mfz8), encoded by the Mfz8 gene (see T89892). It is
 CC an example of a Wnt receptor. Other novel frizzled family members
 CC have been identified in human, mouse and Caenorhabditis elegans
 CC (see W31268-74) and are considered also to be Wnt receptors. Wnt
 CC receptors can be used in a novel, claimed method of screening for
 CC compounds which modulate the binding of a Wnt polypeptide (secreted
 CC proteins involved in cell-to-cell signalling) to a Wnt receptor.
 CC Wnt is involved in (mammary) cancer and other processes involving
 CC growth, development and proliferation (both normal and abnormal).
 CC Modulators identified by the claimed method are useful for
 CC treatment of diseases related to these conditions.
 SQ Sequence 685 AA;

Query Match 19.0%; Score 394; DB 27; Length 685;
 Best Local Similarity 43.8%; Pred. No. 3.38e-28;
 Matches 56; Conservative 23; Mismatches 43; Indels 6; Gaps 6;

Db 40 vplckgigynytmypcnfnhdtqdeaglevhqw-plveicqspdlkfflcsmytpiclc 98
 QY 27 LQCHGIEYQNMRLPNLLGHETMKVEVLEQAGAWIPLVMKQCHPDTKKFLCSLFAFVC 85

Db 99 dykklpprcsvcerakagcaplmrgygfawpdmrdrclpegnpdtlcmdynrtd-lt 157
 QY 86 DLDETIQCHSLCVQVKDRCAPVMSAFGFWPDMLECDREFP-ODN-D-LCPIASSDHL 142

Db 158 taapspr 165
 QY 143 PATEEAPK 150

RESULT 3
 ID W31271 standard; Protein: 585 AA.
 AC W31271:
 DT 27-APR-1998 (first entry)
 DE Human frizzled-5 protein Mfz5 (Wnt receptor).
 KW Wnt receptor; human frizzled-5 protein; Hfz5 gene;
 KW signal transduction; cancer; cell growth; cell proliferation.
 OS Homo sapiens.
 PN W09739357-A1.
 PD 23-OCT-1997.
 PF 11-APR-1997; U06049.
 PR 12-APR-1996; US-015307.
 PA (UYJO) UNIV LELAND STANFORD JUNIOR.

PA (STRD) UNIV LELAND STANFORD JUNIOR.
 PI Andrew D, Bhanot P, Brink M, Hsieh J, Nathans J,
 PI Nusse R, Samos CH, Wangy;
 DR WPI: 97-526631/48.
 DR N-PSDB; T89889.

PT Identification of Wnt receptor binding modulators - useful for
 PT treatment of cancer and growth, development or proliferation related
 PT disorders
 PS Disclosure: Page 37-39; 61pp; English.
 CC This protein comprises the human transmembrane receptor,
 CC frizzled-5 (Hfz5), encoded by the Hfz5 gene (see T89889). It is
 CC an example of a Wnt receptor. Other novel frizzled family members
 CC have been identified in human, mouse and Caenorhabditis elegans
 CC (see W31268-74) and are considered also to be Wnt receptors. Wnt
 CC receptors can be used in a novel, claimed method of screening for
 CC compounds which modulate the binding of a Wnt polypeptide (secreted
 CC proteins involved in cell-to-cell signalling) to a Wnt receptor.
 CC Wnt is involved in (mammary) cancer and other processes involving
 CC growth, development and proliferation (both normal and abnormal).
 CC Modulators identified by the claimed method are useful for
 CC treatment of diseases related to these conditions.
 SQ Sequence 585 AA;

Query Match 18.0%; Score 374; DB 27; Length 585;
 Best Local Similarity 38.8%; Pred. No. 3.66e-26;
 Matches 47; Conservative 29; Mismatches 41; Indels 4; Gaps 4;

Db 20 lvgraaasakpvcqei-t-vpmcrgigynlthmpnqfnhdtqdeaglevhqw-plvei 76
 QY 7 LFGPPDFSKRNSCKRPIPANLQCHGIEYQNMRLPNLLGHETMKVEVLEQAGAWIPLVMK 65

Db 77 qespdlrfllctmytpicldpyhklpprcsvcerakagcspmlrgygfawpdmrscdlr 136
 QY 66 QCHPDPRKFLCSLFAFVCVCLDLDLDTIQCHSLCVQVKDRCAPVMSAFGFWPDMLECDREF 125

Db 137 P 137
 QY 126 P 126

RESULT 4
 ID W31270 standard; Protein: 537 AA.
 AC W31270:
 DT 27-APR-1998 (first entry)
 DE Mouse frizzled-4 protein Mfz4 (Wnt receptor).
 KW Wnt receptor; mouse frizzled-4 protein; Mfz4 gene;
 KW signal transduction; cancer; cell growth; cell proliferation.
 OS Mus musculus.
 PN W09739357-A1.
 PD 23-OCT-1997.
 PF 11-APR-1997; U06049.
 PR 12-APR-1996; US-015307.
 PA (UYJO) UNIV LELAND STANFORD JUNIOR.
 PA (STRD) UNIV LELAND STANFORD JUNIOR.
 PI Andrew D, Bhanot P, Brink M, Hsieh J, Nathans J,
 PI Nusse R, Samos CH, Wangy;
 DR WPI: 97-526631/48.
 DR N-PSDB; T89888.

PT Identification of Wnt receptor binding modulators - useful for
 PT treatment of cancer and growth, development or proliferation related
 PT disorders
 PS Disclosure: Page 34-35; 61pp; English.
 CC This protein comprises the mouse transmembrane receptor,
 CC frizzled-4 (Mfz4), encoded by the Mfz4 gene (see T89888). It is
 CC an example of a Wnt receptor. Other novel frizzled family members
 CC have been identified in human, mouse and Caenorhabditis elegans
 CC (see W31268-74) and are considered also to be Wnt receptors. Wnt
 CC receptors can be used in a novel, claimed method of screening for
 CC compounds which modulate the binding of a Wnt polypeptide (secreted
 CC proteins involved in cell-to-cell signalling) to a Wnt receptor.
 CC Wnt is involved in (mammary) cancer and other processes involving
 CC growth, development and proliferation (both normal and abnormal).
 CC Modulators identified by the claimed method are useful for

DT 27-APR-1998 (first entry)

27 AFK 1998 (first entry)
Mouse frizzled-3 protein Mfz3 (Wnt receptor)

Mouse frizzled-3 protein Mfz3 (Wnt receptor).

KW Wnt receptor; mouse frizzled-3 protein; Mfz3 gene;
 KW signal transduction; cancer; cell growth; cell proliferation.
 OS Mus musculus.
 PN W09739357-A1.
 PD 23-OCT-1997. U06049.
 PF 11-APR-1997; US-015307.
 PR 12-APR-1996; PR-015307.
 PA (UJJO) UNIV JOHNS HOPKINS.
 PA (STRD) UNIV LEFAND STANFORD JUNIOR.
 PI Andrew D, Bhanot P, Brink M, Hsieh J, Nathans J,
 PI Nusse R, Samos CH, Wangy;
 DR WPI; 97-526631/48.
 DR N-PSDB; T89886.
 PT Identification of Wnt receptor binding modulators - useful for
 PT treatment of cancer and growth, development or proliferation related
 PT disorders
 PS Disclosure; Page 27-29; 61pp; English.
 CC This protein comprises the mouse frizzled-3 protein encoded by
 CC the Mfz3 gene (see T89886), and is a Wnt receptor. Other novel
 CC frizzled family members have been identified in human, mouse and
 CC Caenorhabditis elegans (see W31268-74) and are considered also to
 CC be Wnt receptors. Wnt receptors can be used in a novel, claimed
 CC method of screening for compounds which modulate the binding of a
 CC Wnt polypeptide (secreted proteins involved in cell-to-cell
 CC signalling) to a Wnt receptor. Wnt is involved in (mammary) cancer
 CC and other processes involving growth, development and proliferation
 CC (both normal and abnormal). Modulators identified by the claimed
 CC method are useful for treatment of diseases related to these
 CC conditions.
 SQ Sequence 666 AA;
 Query Match 15.3%; Score 317; DB 27; Length 666;
 Best Local Similarity 33.3%; Pred. No. 2,03e-20;
 Matches 44; Conservative 29; Mismatches 54; Indels 5; Gaps 5;
 Db 27 scepi-t-lrmcqdplpyattfmpnlhnydgttaalame-pfhpnmvaldcsrdfrflca 83
 QY 19 NCKPIPANLQCHGIEYQNMRLNGLH-ETMKEVLEQAGAWIPLVMKQCHPDPTKKFLCS 77
 Db 84 lyapicme-ygrvlprrlrcqaycesklmefgypwpedmecsfpdcdepyplrd 142
 QY 78 LFAPVCLDLDLDTIQPCHSLCVQVDRCAPVMSAFGFPWPMLECDRFPQDNDLCIPLAS 137
 Db 143 lnygdptegap 154
 QY 138 SDHLLPATEEAP 149
 RESULT 8
 ID W31272 standard; Protein; 709 AA.
 AC W31272;
 DT 27-APR-1998 (first entry)
 DE Mouse frizzled-6 protein Mfz6 (Wnt receptor).
 KW Wnt receptor; mouse frizzled-6 protein; Mfz6 gene;
 KW signal transduction; cancer; cell growth; cell proliferation.
 OS Mus musculus.
 PN W09739357-A1.
 PD 23-OCT-1997. U06049.
 PF 11-APR-1997; US-015307.
 PR 12-APR-1996; PR-015307.
 PA (UJJO) UNIV JOHNS HOPKINS.
 PA (STRD) UNIV LEFAND STANFORD JUNIOR.
 PI Andrew D, Bhanot P, Brink M, Hsieh J, Nathans J,
 PI Nusse R, Samos CH, Wangy;
 DR WPI; 97-526631/48.
 DR N-PSDB; T89890.
 PT Identification of Wnt receptor binding modulators - useful for
 PT treatment of cancer and growth, development or proliferation related
 PT disorders
 PS Disclosure; Page 41-42; 61pp; English.
 CC This protein comprises the mouse putative transmembrane receptor,
 CC frizzled-6 (Mfz6), encoded by the Mfz6 gene (see T89890). It is
 CC an example of a Wnt receptor. Other novel frizzled family members

CC have been identified in human, mouse and Caenorhabditis elegans
 CC (see W31268-74) and are considered also to be Wnt receptors. Wnt
 CC receptors can be used in a novel, claimed method of screening for
 CC compounds which modulate the binding of a Wnt polypeptide (secreted
 CC proteins involved in cell-to-cell signalling) to a Wnt receptor.
 CC Wnt is involved in (mammary) cancer and other processes involving
 CC growth, development and proliferation (both normal and abnormal).
 CC Modulators identified by the claimed method are useful for
 CC treatment of diseases related to these conditions.
 SQ Sequence 709 AA;
 Query Match 13.7%; Score 284; DB 27; Length 709;
 Best Local Similarity 33.6%; Pred. No. 3.89e-17;
 Matches 41; Conservative 26; Mismatches 49; Indels 6; Gaps 6;
 Db 24 cepitvp-r-cmkmtv-nmtffnlnghydggaavemghflhlanlecsnlemflcqa 80
 QY 20 CKPIPANLQCHGIEYQNMRL-PNLLGHETMKEVLEQAGAWIPLVMKQCHPDPTKKFLCS 78
 Db 81 fipcteqih-vvlprcklcekivsdcklmdtfgirwpeelecnrlphcdtd-vpvtsh 138
 QY 79 FAPVCLDLDLDTIQPCHSLCVQVDRCAPVMSAFGFPWPMLECDRFPQDNDLCIPLAS 138
 Db 139 ph 140
 QY 139 DH 140
 RESULT 9
 ID R74187 standard; Protein; 581 AA.
 AC R74187;
 DT 24-DEC-1995 (first entry)
 DE Chick p75.
 KW Neural axon out-growth modulator; epidermal growth factor; EGF;
 KW netrin-2; p75; neurodegenerative disease; transgenic animal;
 KW gene therapy.
 OS Gallus sp.
 FH Key
 FT peptide
 FT 1..15
 FT /label= Sig_peptide
 FT 16..260
 FT /label= Domain-VI
 FT /note= "Domain VI is probably involved in protein-
 FT protein complex formation"
 FT 261..429
 FT /label= Domain-V
 FT /note= "Domain V includes EGF-like repeats"
 FT 430..581
 FT /note= "c-terminal region"
 PN W09513367-A1.
 PD 18-MAY-1995.
 PF 08-NOV-1994; U12913.
 PR 12-NOV-1993; US-132019.
 PA (UEGC) UNIV CALIFORNIA.
 PA (UJJO) UNIV COLUMBIA NEW YORK.
 PI Dodd J, Jessell T, Kennedy T, Placzek M, Serafini T;
 PI Tessier-Lavigne M;
 DR WPI; 95-194086/25.
 DR N-PSDB; Q92367.
 PT Neural axon out-growth modulators derived from EGF-like repeats of
 PT netrin 1 or netrin 2 - comprise peptide(s) capable of selectively
 PT increasing spinal axon out-growth or directing axon orientation
 PS Claim 1; Page 48-50; 58pp; English.
 CC An E10 chick brain cDNA library was screened with probes based on
 CC netrin-1 (p78) or netrin-2 (p75) sequences to isolate chick p78
 CC and p75 partial cDNA clones. Full-length clones (given in
 CC Q92366-67, respectively) were subsequently obt'd. by 3'RACE. cDNA
 CC is expressed e.g. in COS or insect cells for recombinant p78 and
 CC p75 prodn., used to breed transgenic animals, or for gene therapy.
 SQ Sequence 581 AA;
 Query Match 7.1%; Score 148; DB 14; Length 581;
 Best Local Similarity 27.0%; Pred. No. 3.20e-04;


```
PN W09513367-A1.
PD 18-MAY-1995.
PF 08-NOV-1994; U12913.
PR 12-NOV-1993; US-152019.
PA (REGC ) UNIV CALIFORNIA.
PA (OYCO ) UNIV COLUMBIA NEW YORK.
PI Dodd J, Jessell T, Kennedy T, Placzek M, Serafini T;
PI Tessier-Lavigne M;
DR WPI: 95-194086/25.
DR N-PSDB; Q92368.
PT Neural axon out-growth modulators derived from EGF-like repeats of
PT netrin 1 or netrin 2 - comprise peptide(s) capable of selectively
PT increasing spinal axon out-growth or directing axon orientation
PS Claim 1; Page 51-52; 58pp; English.
CC Chick p75 and p78 sequences (Given in R74186-87, respectively)
CC were used to identify conserved amino acid regions, which were
CC then used to design degenerate primers for the amplification
CC of a fragment of the mouse p78 cDNA (Q92368). This cDNA can
CC be expressed in host cells for recombinant p78 prodn, or used
CC to breed transgenic animals, or for gene therapy.
SQ Sequence 529 AA;

Query Match 5.1%; Score 105; DB 14; Length 529;
Best Local Similarity 23.5%; Pred. No. 1.50e+00;
Matches 24; Conservative 25; Mismatches 45; Indels 8; Gaps 7;

Db 388 svxexedcdsyckaskgklmmkkyckrdyavqihilkadkagdw-wkftvnlls-- 444
QY 143 PATEAPKVCBA-CNNKDDNDIMEILCKNDFALKTKV-KEITYINRDKIILETKSKT 200
Db 445 vykqgtsrtrgdgs-lwrsrdiackcpkplkk-yllylg 484
QY 201 IYKLVGVSERDLKKSVLWKD-SLQCTCEENNDINAPYLVMG 241

RESULT 13
ID W26642 standard; Protein; 732 AA.
AC W26642;
DT 11-FEB-1998 (first entry)
DE Human RECK cancer-inhibiting protein.
KW RECK; reversion-inducing cysteine rich protein with Kazal motif;
KW human; cancer; gene therapy; diagnosis.
OS Homo sapiens.
FH Key Location/Qualifiers
FT Domain 396..415
FT /label= Kazal_domain
PN W09724439-A1.
PD 10-JUL-1997.
PF 24-DEC-1996; U20812.
PR 27-DEC-1995; JP-340469.
PA (AMGE-) AMGEN INC.
PA (KITA/) KITAYAMA H.
PA (NODA/) NODA M.
PA (TAKA/) TAKAHASHI C.
PA (SANY ) SANKYO CO LTD.
PI Kitayama H, Noda M, Takahashi C;
DR WPI: 97-363675/33.
DR N-PSDB; T90508.
PT RECK gene and corresponding protein sequences - enables reversion of
PT cancer cells
PS Claim 7; Page 42-44; 53pp; English.
CC This protein sequence comprises the human RECK protein (reversion-
CC inducing cysteine rich protein with Kazal motif), which is capable
CC of transforming malignant cancer cells with an activated ras gene
CC into normal cells (reversion activity). Its amino acid sequence
CC was deduced from a cDNA clone (see T90508) obtained from human
CC fibroblast MRC-5 (ATCC CCL 171) cells. A claimed method for
CC treatment of cancer comprises contacting the cancer cells with a
CC RECK polypeptide. RECK may also be expressed using gene therapy
CC methods for in vivo treatment of cancer.
SQ Sequence 732 AA;

Query Match 4.8%; Score 100; DB 25; Length 732;

PN W09513367-A1.
PD 18-MAY-1995.
PF 08-NOV-1994; U12913.
PR 12-NOV-1993; US-152019.
PA (REGC ) UNIV CALIFORNIA.
PA (OYCO ) UNIV COLUMBIA NEW YORK.
PI Dodd J, Jessell T, Kennedy T, Placzek M, Serafini T;
PI Tessier-Lavigne M;
DR WPI: 95-194086/25.
DR N-PSDB; Q92368.
PT Neural axon out-growth modulators derived from EGF-like repeats of
PT netrin 1 or netrin 2 - comprise peptide(s) capable of selectively
PT increasing spinal axon out-growth or directing axon orientation
PS Claim 1; Page 51-52; 58pp; English.
CC Chick p75 and p78 sequences (Given in R74186-87, respectively)
CC were used to identify conserved amino acid regions, which were
CC then used to design degenerate primers for the amplification
CC of a fragment of the mouse p78 cDNA (Q92368). This cDNA can
CC be expressed in host cells for recombinant p78 prodn, or used
CC to breed transgenic animals, or for gene therapy.
SQ Sequence 529 AA;

Query Match 5.1%; Score 105; DB 14; Length 529;
Best Local Similarity 23.5%; Pred. No. 1.50e+00;
Matches 24; Conservative 25; Mismatches 45; Indels 8; Gaps 7;

Db 388 svxexedcdsyckaskgklmmkkyckrdyavqihilkadkagdw-wkftvnlls-- 444
QY 143 PATEAPKVCBA-CNNKDDNDIMEILCKNDFALKTKV-KEITYINRDKIILETKSKT 200
Db 445 vykqgtsrtrgdgs-lwrsrdiackcpkplkk-yllylg 484
QY 201 IYKLVGVSERDLKKSVLWKD-SLQCTCEENNDINAPYLVMG 241

RESULT 13
ID W26642 standard; Protein; 732 AA.
AC W26642;
DT 11-FEB-1998 (first entry)
DE Human RECK cancer-inhibiting protein.
KW RECK; reversion-inducing cysteine rich protein with Kazal motif;
KW human; cancer; gene therapy; diagnosis.
OS Homo sapiens.
FH Key Location/Qualifiers
FT Domain 396..415
FT /label= Kazal_domain
PN W09724439-A1.
PD 10-JUL-1997.
PF 24-DEC-1996; U20812.
PR 27-DEC-1995; JP-340469.
PA (AMGE-) AMGEN INC.
PA (KITA/) KITAYAMA H.
PA (NODA/) NODA M.
PA (TAKA/) TAKAHASHI C.
PA (SANY ) SANKYO CO LTD.
PI Kitayama H, Noda M, Takahashi C;
DR WPI: 97-363675/33.
DR N-PSDB; T90508.
PT RECK gene and corresponding protein sequences - enables reversion of
PT cancer cells
PS Claim 7; Page 42-44; 53pp; English.
CC This protein sequence comprises the human RECK protein (reversion-
CC inducing cysteine rich protein with Kazal motif), which is capable
CC of transforming malignant cancer cells with an activated ras gene
CC into normal cells (reversion activity). Its amino acid sequence
CC was deduced from a cDNA clone (see T90508) obtained from human
CC fibroblast MRC-5 (ATCC CCL 171) cells. A claimed method for
CC treatment of cancer comprises contacting the cancer cells with a
CC RECK polypeptide. RECK may also be expressed using gene therapy
CC methods for in vivo treatment of cancer.
SQ Sequence 732 AA;

Query Match 4.8%; Score 100; DB 25; Length 732;
```

```
Best Local Similarity 30.5%; Pred. No. 3.78e+00;
Matches 25; Conservative 15; Mismatches 36; Indels 6; Gaps 5;

Db 158 ipvldikkqpemwkaiaclqikpchkshksrgsli-cksdceveilkkgd-qnkpfedht 215
QY 60 IP-LVMKQCHPDT-KKFLCSLFAVCLDLDLDETQPCSLCVQVKDRCAPVMSAFGFWP 117
Db 216 aesicellsgtdldkncipldt 237
QY 118 DMLECDRFFQDNDL--CIPLAS 137

RESULT 14
ID W55368 standard; Protein; 804 AA.
AC W55368;
DT 17-JUN-1998 (first entry)
DE H. pylori ORF 09apil406orf2 protein.
KW Cytoplasmic; vaccine; prevention; treatment; infection; envelope;
KW identification; binding compound; bacteria; life cycle; activator;
KW inhibitor; duodenal ulcer disease; chronic gastritis; diagnosis.
OS Helicobacter pylori.
PN W09737044-A1.
PD 09-OCT-1997.
PF 27-MAR-1997; U05223.
PR 06-DEC-1996; US-761318.
PR 29-MAR-1996; US-625811.
PR 02-APR-1996; US-758731.
PR 25-OCT-1996; US-736905.
PR 28-OCT-1996; US-738859.
PA (ASTR ) ASTRA AB.
PI Alm RA, Smith D;
DR WPI: 97-503122/46.
DR N-PSDB; V24777.
PT Helicobacter pylori nucleic acid sequences and encoded
PT polypeptide(s) - useful in vaccines to treat or prevent H. pylori
PT infection and for diagnosis of H. pylori infection.
PS Claim 14; Page 587-589; 1145pp; English.
CC This sequence is a H. pylori protein of unspecified function.
CC The protein may be used in a vaccine to prevent or treat H. pylori
CC infection or to identify H. pylori polypeptide binding compounds. The
CC useful as potential H. pylori life cycle activators or inhibitors. The
CC DNA and probes derived from it may be used for the identification of
CC H. pylori in a sample and the diagnosis of H. pylori infection. Nucleic
CC acid sequences complementary to the DNA act as antisense sequences and
CC can be used to prevent the translation of H. pylori mRNA. Antibodies
CC against the protein can be used in immunoassays to evaluate the abundance
CC and distribution of H. pylori-specific antigens. The genomic sequence of
CC H. pylori (ATCC 35679) was determined from overlapping contigs generated
CC by mechanically shearing the bacterial DNA. The sequences were analysed
CC for ORF of at least 180 nucleotides, and the predicted coding regions
CC defined by computer evaluation. To identify likely H. pylori antigens for
CC vaccine development, the amino acid sequences predicted from various ORF
CC were analysed for significant homology to other known or exported
CC membrane proteins. Having identified and determined the sequences of
CC interest, particular regions can be isolated from H. pylori by PCR
CC amplification for recombinant polypeptide production, e.g. in E. coli
CC hosts.
SQ Sequence 804 AA;

Query Match 4.8%; Score 99; DB 29; Length 804;
Best Local Similarity 31.5%; Pred. No. 4.53e+00;
Matches 17; Conservative 16; Mismatches 19; Indels 2; Gaps 2;

Db 78 vidsipkegftplenafnitysmrqggfilhikpkmrlltfsdrdykai 131
QY 165 IMETLCKNDFA-LKIKVKEITINRDKIILETKSKTIYKLVGVS-ERDLKRSV 216

RESULT 15
ID R44929 standard; Protein; 15281 AA.
AC R44929;
DT 08-JUL-1994 (first entry)
DE T. niveum Cyclosporin synthetase.
```


KW Enzyme; cyclosporin; synthetase-like activity; Tolypocladium niveum;
KW T. inflatum GMS; biosynthesis; vector; cyclosporin synthetase.
OS Tolypocladium niveum.
PN EP-578616-A.
PD 12-JAN-1994.

PF 05-JUL-1993; 810474.

PR 09-JUL-1992; AT-001403.

PR 08-MAR-1993; AT-000437.

PR 29-APR-1993; CH-001310.

PR 04-MAY-1993; CH-001375.

PA (SANO) SANDOZ LTD.

PA (SANO) SANDOZ PATENT GMBH.

PA (SANO) SANDOZ-ERFINDUNGEN VERW GES MBH.

PI Leitner E, Schneider E, Schoengendorfer K, Weber G.

DR WPI: 94-010432/02.

DR N-PSDB: Q54385.

PT Isolated DNA sequence - which codes for enzyme having cyclosporin

PT synthetase like activity

PS Claim 1; Page 41-84; 93pp; English.

CC This sequence represents an enzyme which has cyclosporin synthetase-

CC like activity. This sequence was isolated from Tolypocladium niveum

CC (formerly known as T. inflatum GMS). This enzyme catalyses the

CC peptide biosynthesis of cyclosporins and structurally related

CC molecules. This sequence may be used for the production of

CC cyclosporin by transforming a vector containing this sequence in

CC to a recombinant host. This allows effective production of anti-

CC biotic cyclosporin or its derivatives.

SQ Sequence 15281 AA;

Query Match 4.7%; Score 97; DB 9; Length 15281;

Best Local Similarity 26.7%; Pred. No. 6.52e+00;

Matches 24; Conservative 23; Mismatches 35; Indels 8; Gaps 8;

Db 3819 lrgtkisdhiaianipnsktiverticesvydlgg-dakdsndrswlsaa-rsnakva 3876

QY 176 LK-IKVKE-ITVIN-RDTRILE-TKSKTIYKLVGVSRDLKSLWLKDSLQCTCEMN 231

Db 3877 slsaldldiaaeagfrveiscarqwsqg 3906

QY 232 DINAPYLW-MGOKQGGELVITSVKRW-QKG 259

Search completed: Tue Aug 4 10:40:20 1998

Job time : 71 secs.

THIS PAGE BLANK (USPTO)

W P S R L
***** (TW)

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPsrch_pp protein - protein database search, using Smith-Waterman algorithm
Run on: Tue Aug 4 10:40:37 1998; MasPar time 15.25 Seconds
Tabular output not generated.

Title: >US-08-848-439-3
Description: (1-275) from US08848439.pep
Perfect Score: 2073
Sequence: 1 SARGFLFGQPDFSYKRSNC.....WQKQREFKRISIRKLOK 275

Scoring table: PAM 150
Gap 11

Searched: 120441 seqs, 36531193 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: pir56
1:pir1 2:pir2 3:pir3 4:pir4 5:nrl3d

Statistics: Mean 44.556; Variance 87.725; scale 0.508

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result		Query		Match		Length		DB		ID		Description		Pred. No.	
1	386	18	6	641	2	A45054	2	probable intercellular	1.05e-51						
2	355	17	1	694	2	S78444	2	dfz2 protein - fruit	1.19e-45						
3	355	17	1	694	2	S71786	2	dfz2 protein - fruit	1.19e-45						
4	332	16	0	415	2	S15709	2	hypothetical protein	3.34e-41						
5	332	16	0	581	2	S03540	2	gene frizzled protein	3.34e-41						
6	178	8	6	562	2	B56101	2	collagen alpha 1(XVII	4.96e-13						
7	148	7	1	581	2	B54665	2	netrin-2 precursor -	4.04e-08						
8	124	6	0	606	2	A54665	2	netrin-1 precursor -	1.87e-04						
9	122	5	9	793	2	JC5539	2	Smoothed protein	3.86e-04						
10	101	4	9	1735	2	S22812	2	DNA-directed RNA poly	2.94e-01						
11	99	4	8	322	2	H64825	2	hypothetical protein	5.34e-01						
12	100	4	8	500	2	D31579	2	cytochrome P450 2D4 -	3.97e-01						
13	100	4	8	500	2	I52313	2	cytochrome P450 2D4 h	3.97e-01						
14	99	4	8	724	2	B35711	2	ribosomal protein S6	5.34e-01						
15	98	4	7	391	2	E64474	2	hypothetical protein	7.17e-01						
16	98	4	7	1032	1	G7BPT4	1	baseplate protein gp7	7.17e-01						
17	97	4	7	15281	2	S41309	2	cyclosporin synthetas	9.62e-01						
18	96	4	6	520	2	J50291	2	intermediate filament	1.29e+00						
19	96	4	6	735	2	I51901	2	ribosomal protein S6	1.29e+00						
20	94	4	5	271	2	E64694	2	hypothetical protein	2.29e+00						
21	93	4	5	735	2	A53300	2	ribosomal protein S6	3.04e+00						
22	92	4	4	203	2	B70254	2	conserved hypothetical	4.03e+00						
23	92	4	4	608	2	B70188	2	hypothetical protein	4.03e+00						

24 92 4.4 801 2 B64679 paralysed flagella pr 4.03e+00
25 92 4.4 1562 2 S53059 probable membrane pro 4.03e+00
26 90 4.3 176 2 S70008 troponin I - Atlantic 7.04e+00
27 89 4.3 262 2 A60105 virF protein - Shigel 9.27e+00
28 89 4.3 262 2 A47605 virF virulence protei 9.27e+00
29 89 4.3 262 2 S14646 virF protein - Shigel 9.27e+00
30 90 4.3 263 2 S73489 S-adenosylmethionine- 7.04e+00
31 89 4.3 508 2 A64571 GMP synthase - Helico 9.27e+00
32 90 4.3 521 2 A53153 glucose transport pro 7.04e+00
33 89 4.3 752 2 A32571 ribosomal protein S6 9.27e+00
34 88 4.2 127 2 S45967 probable membrane pro 1.22e+01
35 88 4.2 182 2 S28310 hypothetical protein 1.22e+01
36 88 4.2 183 2 D69108 phyocyanin alpha phy 1.22e+01
37 88 4.2 370 2 A69430 conserved hypothetical 1.22e+01
38 88 4.2 443 2 JQ1527 alpha-amylase (EC 3.2 1.22e+01
39 88 4.2 445 2 S19990 alpha-amylase (EC 3.2 1.22e+01
40 88 4.2 633 2 C32571 ribosomal protein S6 1.22e+01
41 88 4.2 740 2 I38556 insulin-stimulated pr 1.22e+01
42 88 4.2 849 2 S00030 neurofilament triplet 1.22e+01
43 88 4.2 871 2 I48696 gene Nsk2 protein - m 1.22e+01
44 88 4.2 881 2 I48697 gene Nsk2 protein - m 1.22e+01
45 88 4.2 1102 2 S44772 C29E4.4 protein - Cae 1.22e+01

ALIGNMENTS

1
RESULT #type complete
ENTRY probable intercellular signal transducer or transmitter Fz-1
TITLE - rat
ORGANISM #formal_name Rattus norvegicus #common_name Norway rat
DATE 30-Apr-1993 #sequence_revision 18-Nov-1994 #text_change 28-Apr-1995
ACCESSIONS A45054
REFERENCE A45054
#authors Chan, S.D.; Karpf, D.B.; Fowlkes, M.E.; Hooks, M.; Bradley, M.S.; Vuong, V.; Bambino, T.; Liu, M.Y.; Arnaud, C.D.; Strewler, G.J.; Nissenson, R.A.
#journal J. Biol. Chem. (1992) 267:25202-25207
#title Two homologs of the Drosophila polarity gene frizzled (fz) are widely expressed in mammalian tissues.
#cross-references MIM:39094226
#accession A45054
#status preliminary; not compared with conceptual translation
#molecule_type nucleic acid
#residues 1-641 #label CHA
#experimental_source UMR 106 osteosarcoma cell line
#note sequence extracted from NCBI backbone (NCBIP:120154) #length 641 #molecular-weight 71054 #checksum 8376

SUMMARY
Query Match 18.6%; Score 386; DB 2; Length 641;
Best Local Similarity 39.2%; Pred. No. 1.05e-51;
Matches 49; Conservative 27; Mismatches 43; Indels 6; Gaps 6;
Db 108 HGYCOPI-S-IPLTDTAYNOTIPNLLGHTNQDAGLEVHOFYPLVKVQCSAEKFLC 165
Qy 17 RSNCKPIANLQCHGIEYQNMRLPLGHTMKTEVLEQAGWIPYMKQCHPDTKKFLC 76
Db 166 SMYAPVC-TVLEQALPCRSICERAOQ-CEALMKNKFGOWPDTLKCEKFPVHGRGELCVG 223
Qy 77 SLFAPVCLDDUDETQCHSLCVQVKRCAPVMSAFGFPNPDMLCEDRFP-QD-NDLCIP 134
Db 224 QNTSD 228
Qy 135 LASSD 139

2
RESULT #type complete
ENTRY dfz2 protein - fruit fly (Drosophila melanogaster)
TITLE #formal_name Drosophila melanogaster
ORGANISM 12-Feb-1998 #sequence_revision 13-Mar-1998 #text_change 13-Mar-1998

ACCESSIONS S78444
REFERENCE S78444
#authors Bhanot, P.; Wang, Y.; Nathans, J.
#journal Submitted to the EMBL Data Library, July 1996
#title S78444
#accession S78444
#status preliminary
#molecule_type DNA
#residues 1-694 #label BHA
#cross-references EMBL:U65589
#length 694 #molecular-weight 75437 #checksum 3113

Query Match 17.1%; Score 355; DB 2; Length 694;
Best Local Similarity 36.2%; Pred. No. 1.19e-45;
Matches 47; Conservative 32; Mismatches 45; Indels 6; Gaps 6;
Db 51 YGVPAIP-KDNLRCETITPMCRGIGYNTSPNEMNHETQDEAGLEVHGFQW-PLVEIK 108
QY 8 FGQPDFSYKRSNCKPPIANLQCHGIEYQNNRPLNLGHETMKREV-LEQAGAWIPLVMKQ 66
Db 109 CSPDLKFFLCSMTPIGLEDYHKPLPVCRSVCERARSGCAPIMQOYSFEPWERMACEHLP 168
QY 67 CHPTDKFLCSLFPVCLDLDLDTIQCHSLCVQKDRCAPVMSAFGFPWPDMLCEDRFP 126
Db 169 LHGDPDNLCM 178
QY 127 -Q-D-NDLCI 133

RESULT 3
ENTRY dfz2 protein - fruit fly (Drosophila melanogaster)
TITLE
ORGANISM
DATE 12-Feb-1998 #sequence_revision 13-Mar-1998 #text_change
13-Mar-1998
ACCESSIONS S71786
REFERENCE S71786
#authors Bhanot, P.; Brink, M.; Harryman Samos, C.; Hsieh, J.C.; Wang,
Y.; Macke, J.P.; Andrew, D.; Nathans, J.; Nusse, R.
#journal Nature (1996) 382:225-230
#title A new member of the frizzled family from Drosophila functions
as a wingless receptor.
#accession S71786
#status preliminary; nucleic acid sequence not shown
#molecule_type DNA
#residues 1-694 #label BHA
#cross-references EMBL:U65589
#length 694 #molecular-weight 75423 #checksum 3095

Query Match 17.1%; Score 355; DB 2; Length 694;
Best Local Similarity 36.2%; Pred. No. 1.19e-45;
Matches 47; Conservative 32; Mismatches 45; Indels 6; Gaps 6;
Db 51 YGVPAIP-KDNLRCETITPMCRGIGYNTSPNEMNHETQDEAGLEVHGFQW-PLVEIK 108
QY 8 FGQPDFSYKRSNCKPPIANLQCHGIEYQNNRPLNLGHETMKREV-LEQAGAWIPLVMKQ 66
Db 109 CSPDLKFFLCSMTPIGLEDYHKPLPVCRSVCERARSGCAPIMQOYSFEPWERMACEHLP 168
QY 67 CHPTDKFLCSLFPVCLDLDLDTIQCHSLCVQKDRCAPVMSAFGFPWPDMLCEDRFP 126
Db 169 LHGDPDNLCM 178
QY 127 -Q-D-NDLCI 133

RESULT 4
ENTRY hypothetical protein - fruit fly (Drosophila melanogaster)
TITLE
ORGANISM
DATE 21-Nov-1993 #sequence_revision 10-Nov-1995 #text_change
16-Feb-1997
ACCESSIONS S15709
REFERENCE S15709

#authors Adler, P.N.; Vinson, C.; Park, W.J.; Conover, S.; Klein, L.
#journal Genetics (1990) 126:401-416
#title Molecular structure of frizzled, a Drosophila tissue polarity
gene.
#cross-references MUID:91060073
#accession S15709
#status preliminary
#molecule_type DNA
#residues 1-415 #label ADL
#cross-references EMBL:X54648
GENETICS
#gene FlyBase:fz
#cross-references FlyBase:FBgn0001085
#introns 224/3; 264/3; 329/3; 405/1
#length 415 #molecular-weight 46117 #checksum 4401

Query Match 16.0%; Score 332; DB 2; Length 415;
Best Local Similarity 37.5%; Pred. No. 3.34e-41;
Matches 45; Conservative 22; Mismatches 47; Indels 6; Gaps 6;
Db 50 HNRCEPI-T-ISICKNIPYNTIMPNLIGHTKQEEAGLEVHGFAPLVKIGSCDDQLQLFLC 107
QY 17 RSNCKPIANLQCHGIEYQNNRPLNLGHETMKREVLEQAGAWIPLVMKQCHPDTKKFLC 76
Db 108 SLYVPVC-TILERPIPPCRSLCESARV-CERLMKTYNFNPNLENLSCSKPVPVHGEGDLVYA 165
QY 77 SLFAPVCLDLDLDTIQCHSLCVQKDRCAPVMSAFGFPWPDMLCEDRFP-QD-NDLCIP 134

RESULT 5
ENTRY S03540 #type complete
TITLE gene frizzled protein precursor - fruit fly (Drosophila melanogaster)
ORGANISM #formal_name Drosophila melanogaster
DATE 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change
23-Aug-1997
ACCESSIONS S03540; S15708
REFERENCE S03540
#authors Vinson, C.R.; Conover, S.; Adler, P.N.
#journal Nature (1989) 338:263-264
#title A Drosophila tissue polarity locus encodes a protein
containing seven potential transmembrane domains.

#cross-references MUID:89159415
#accession S03540 not compared with conceptual translation
#molecule_type mRNA
#residues 1-581 #label VIN
REFERENCE S15708
#authors Adler, P.N.; Vinson, C.; Park, W.J.; Conover, S.; Klein, L.
#journal Genetics (1990) 126:401-416
#title Molecular structure of frizzled, a Drosophila tissue polarity
gene.
#cross-references MUID:91060073
#accession S15708
#status preliminary
#molecule_type DNA
#residues 1-581 #label ADL
#cross-references EMBL:X54648

GENETICS
#gene FlyBase:fz
#cross-references FlyBase:FBgn0001085
#introns 224/3; 264/3; 329/3; 405/1
#length 581 #molecular-weight 64846 #checksum 358
#domain signal sequence #status predicted #label SIG\

KEYWORDS
FEATURE 1-26
27-581
#product gene frizzled protein #status predicted #label
#text_change
#length 581 #molecular-weight 64846 #checksum 358
Query Match 16.0%; Score 332; DB 2; Length 581;
Best Local Similarity 37.5%; Pred. No. 3.34e-41;
Matches 45; Conservative 22; Mismatches 47; Indels 6; Gaps 6;

```
QY      133 IPLASDHLLPATEAPKVCEA-CKNKNDNDINMETLCKNDFALKIKVKEITYINRDTK 191
Db      490 FTNIIS--VYKORDERVKR-GDNFWIHLKOLSCPKXI-QISKXYLVNG 536
        : :: : ||| : : ||| : | | : : | : | | | | |
QY      192 ILLETSKTIYKUNGVSERDLKKSJWLK-DSLQCTCEENMDINAPYLVNG 241
```

```

RESULT      8
ENTRY
TITLE       A54665      #type complete
ENTRY      netrin-1 precursor - chicken
TITLE      #formal_name Gallus gallus #common_name chicken
ORGANISM   06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change
DATE       10-Sep-1997
ACCESSIONS A54665
REFERENCE   A54665
AUTHORS    Serafini, T.; Kennedy, T.E.; Galko, M.J.; Mirzayan, C.;
            Jessell, T.M.; Tessier-Lavigne, M.
JOURNAL    Cell (1994) 78:409-424
TITLE      The netrins define a family of axon outgrowth-promoting
            proteins homologous to Caenorhabditis elegans UNC-6.
ACCESSION A54665
STATUS     preliminary
MOLECULE   molecule_type mRNA
RESIDUES   1-606    #label SER
CROSS-REF  #cross-references GB:I34549; NID:g529418;
            #length 606 #molecular_weight 68126 #checksum 5400
SUMMARY

```

Query Match 6.0%; Score 124; DB 2; Length 606;
Best Local Similarity 26.1%; Pred. No. 1.87e-04;
Matches 31: Conservative 34; Mismatches 41; Indels 13; Gaps 11;

```
Db      456 IPAAPPTAASSTEEPAD-CDSYCKASKGLKINMKKYCKDYAVQIHLKAENADW-W 513
        ||||| :||||: |::||| : | ||:|::| : | ::
Ov     133 IPIASSDHLIPATEAPKVGCA-CKNKNDNDDIMETLCKNDFAKKIV-KEITYINRDT 190
```

RESULT	9
ENTRY	JC5539
TITLE	Smoothered protein - mouse
ORGANISM	#formal_name Mus musculus #common_name house mouse
DATE	03-Sep-1997 #sequence_revision 05-Sep-1997 #text_change 05-Sep-1997
ACCESSIONS	JC5539; PC4476

ACCESSIONS	JC5539	Akiyama, H.; Shigeno, C.; Hiraki, Y.; Shukunami, C.; Kohno, H.; Akagi, M.; Konishi, J.; Nakamura, T.
REFERENCE	JC5539	Biotech. Biochem. Commun. (1997) 235:142-147
#authors	JC5539	Cloning of a mouse Smoothed cDNA and expression patterns of hedgehog signalling molecules during chondrogenesis and cartilage differentiation in clonal mouse EC cells, ATDC5.
#journal	JC5539	
#title	JC5539	
#accession	JC5539	

```

#accession JC5539
##molecule_type mRNA
##residues 1-793 ##label AKI
#accession PC4476
##molecule_type protein
##residues 528-533;539-545;600-605 ##label AK2
##experimental_source ADTC5 cell
#COMMENT
This protein is used in the conserved targets in Hedgehog
signalling pathway, together with Patched and Gli. These proteins
are responsible for the skeletal abnormalities in Gorlin and
Greig syndromes.

```

```
FEATURE
1-32
SUMMARY
#domain signal sequence #status predicted #label SIG
#length 793 #molecular-weight 87299 #checksum 8609
```

Query Match 5.98; Score 122; DB 2; Length 793;
Best Local Similarity 26.78; Pred. No. 3,66e-04;
Matches 31; Conservative 22; Mismatches 54; Indels 9; Gaps 7;

```

##cross-references GB:A5000189; GB:U00096; NID:g1787097; PID:g1787098.8
OWGP:b0872
##experimental_source strain K-12, substrain MG1655
SUMMARY #length 322 #molecular-weight 35740 #checksum 5190

Query Match 4.88; Score 99; DB 2; Length 322;
Best Local Similarity 27.59; Pred. No. 5.34e-01;
Matches 19; Conservative 22; Mismatches 21; Indels 7; Gaps 7;

Db 87 RDVKRGDYLWLSDAMGFTCCDKAE-DK-FLLLAA-GCGVTPIMSRRLAKNRPOADV R 143
QY 210 RDLKKS-V-LWLKDSL-QCTCEEMNDINAPLYMGQGGELVITSVKRW-QKGQREFK-R 265
Db 144 VIXNVRTPP 152
QY 266 ISRSIRKLQ 274

RESULT 12
ENTRY D31579 #type complete
TITLE cytochrome P450 2D4 - rat
ALTERNATE_NAMES cytochrome P450CMF3
ORGANISM #formal_name Rattus norvegicus #common_name Norway rat
DATE 28-Feb-1990 #sequence_revision 17-May-1996 #text_change
23-Jan-1998
ACCESSIONS S16873; D31579
REFERENCE S16871
#authors Matsunaga, E.; Umeno, M.; Gonzalez, F.J.
#journal J. Mol. Evol. (1990) 30:155-169
#title The rat P450 IID subfamily: complete sequences of four
closely linked genes and evidence that gene conversions
maintained sequence homogeneity at the heme-binding region
of the cytochrome P450 active site.
#cross-references MUID:90189185
#accession S16873
#molecule_type DNA
#residues 1-500 #label MAT
#cross-references EMBL:X52029; NID:g57815; PID:g57816
REFERENCE A90151
#authors Ishida, N.; Tawaragi, Y.; Inuzuka, C.; Sugita, O.; Kubota,
I.; Nakazato, H.; Noguchi, T.; Sassa, S.
#journal Biochem. Biophys. Res. Commun. (1988) 156:681-688
#title Four species of cDNAs for cytochrome P450 isozymes
immunorelated to P450c-M/F encode for members of P450IID
subfamily, increasing the number of members within the
subfamily.
#cross-references MUID:89050091
#accession D31579
#molecule_type mRNA
#residues 177-500 #label ISH
#cross-references EMBL:M22331; NID:g203829; PID:g203830
GENETICS
#gene CYP2D4
#introns 63/3; 121/1; 172/1; 225/3; 284/3; 332/1; 394/3; 442/1
CLASSIFICATION #superfamily cytochrome P450
KEYWORDS chromoprotein; electron transfer; endoplasmic reticulum;
heme; iron; monooxygenase; oxidoreductase; transmembrane
protein
FEATURE
446 #binding_site heme iron (Cys) (axial ligand) #status
predicted
SUMMARY #length 500 #molecular-weight 56697 #checksum 7863

Query Match 4.88; Score 100; DB 2; Length 500;
Best Local Similarity 26.78; Pred. No. 3.97e-01;
Matches 16; Conservative 22; Mismatches 18; Indels 4; Gaps 3;

Db 173 FPS-PNTLLDKAVCNVIAISLLFACR-FEYNDPFRRL--DLLKDTLDEESGFLPMLN 228
QY 6 FLFGQPDFSKRSNCRPIANLQICGIEFQNNRLPNLIGHETMKVEVLQAGAWIPLVVK 65

RESULT 13

```

```
ENTRY      I52313      #type complete
TITLE      cytochrome P450 2B4 homolog, brain (clone 2d-29) - rat
ORGANISM   #formal_name Rattus norvegicus #common_name Norway rat
DATE       27-Feb-1997 #sequence_revision 27-Feb-1997 #text_change
23-Jan-1998
ACCESSIONS I52313
REFERENCE   I52313
#authors   Kawashima, H.; Strobel, H.W.
#journal   Biochem. Biophys. Res. Commun. (1995) 209:535-540
#title     cDNA cloning of a novel rat brain cytochrome P450 belonging
           to the CYP2D subfamily.
#cross-references MUID:95251650
#accession  I52313
#status    preliminary; translated from GB/EMBL/DBJ
#molecule_type mRNA
#residues  1-500 #label RES
#cross-references GB:S77859; NID:g998524; PID:g998525
#experimental_source brain, strain Sprague-Dawley
CLASSIFICATION #superfamily cytochrome P450
KEYWORDS      heme; transmembrane protein
FEATURE       446
             #binding_site heme iron (Cys) (axial ligand) #status
             predicted
SUMMARY       #length 500 #molecular-weight 56683 #checksum 8540
Query Match   4.8%; Score 99; DB 2; Length 724;
Best Local Similarity 35.2%; Pred. No. 5.34e-01;
Matches 19; Conservative 13; Mismatches 18; Indels 4; Gaps 4;

Db 499 KI-LRQKFFSEAEFVLHTISKVEYLHSGVQVHRLDKPSNLYVDESGNPEC 541
|| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 177 KIKVKEITYINRDTKILLETSKTI-Y-KLNGVSRDLKKS-VLWIKDSLQCTC 227

RESULT      15
ENTRY      60-310
TITLE      #domain protein kinase homology #label KIN1\
68-76      #region protein kinase ATP-binding motif\
405-664    #domain protein kinase homology #label KIN2
SUMMARY    #length 724 #molecular-weight 81594 #checksum 4206;

Query Match   4.8%; Score 99; DB 2; Length 724;
Best Local Similarity 35.2%; Pred. No. 5.34e-01;
Matches 19; Conservative 13; Mismatches 18; Indels 4; Gaps 4;

Db 499 KI-LRQKFFSEAEFVLHTISKVEYLHSGVQVHRLDKPSNLYVDESGNPEC 541
|| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 177 KIKVKEITYINRDTKILLETSKTI-Y-KLNGVSRDLKKS-VLWIKDSLQCTC 227

RESULT      14
ENTRY      B32571
TITLE      #type complete
ORGANISM   ribosomal protein S6 kinase II (EC 2.7.-.-) alpha chain
           homolog (clone Mu6A) - mouse
#formal_name Mus musculus #common_name house mouse
DATE       31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change
26-Feb-1998
ACCESSIONS B32571
REFERENCE   A93113
#authors   Alcorta, D.A.; Crews, C.M.; Sweet, L.J.; Bankston, L.; Jones,
           S.W.; Erikson, R.L.
#journal   Mol. Cell. Biol. (1989) 9:3850-3859
#title     Sequence and expression of chicken and mouse rsk: homologs of
           Xenopus laevis ribosomal S6 kinase.
#cross-references MUID:89384612
#accession  B32571
#molecule_type mRNA
#residues  1-724 #label ALC
#cross-references GB:M23489
CLASSIFICATION #superfamily ribosomal protein S6 kinase II; protein kinase
           homology
KEYWORDS      ATP; phosphotransferase; serine/threonine-specific protein
           kinase
FEATURE       60-310
           #domain protein kinase homology #label KIN1\
68-76      #region protein kinase ATP-binding motif\
405-664    #domain protein kinase homology #label KIN2
SUMMARY    #length 724 #molecular-weight 81594 #checksum 4206;

Query Match   4.8%; Score 99; DB 2; Length 724;
Best Local Similarity 35.2%; Pred. No. 5.34e-01;
Matches 19; Conservative 13; Mismatches 18; Indels 4; Gaps 4;

Db 499 KI-LRQKFFSEAEFVLHTISKVEYLHSGVQVHRLDKPSNLYVDESGNPEC 541
|| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 177 KIKVKEITYINRDTKILLETSKTI-Y-KLNGVSRDLKKS-VLWIKDSLQCTC 227

RESULT      15
ENTRY      60-310
TITLE      #domain protein kinase homology #label KIN1\
68-76      #region protein kinase ATP-binding motif\
405-664    #domain protein kinase homology #label KIN2
SUMMARY    #length 724 #molecular-weight 81594 #checksum 4206;

Query Match   4.8%; Score 99; DB 2; Length 724;
Best Local Similarity 35.2%; Pred. No. 5.34e-01;
Matches 19; Conservative 13; Mismatches 18; Indels 4; Gaps 4;

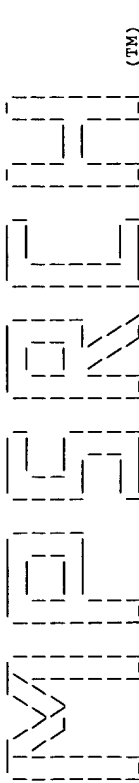
Db 499 KI-LRQKFFSEAEFVLHTISKVEYLHSGVQVHRLDKPSNLYVDESGNPEC 541
|| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 177 KIKVKEITYINRDTKILLETSKTI-Y-KLNGVSRDLKKS-VLWIKDSLQCTC 227
```

```
ENTRY      E64474      #type complete
TITLE      hypothetical protein MJ1398 - Methanococcus jannaschii
ORGANISM   #formal_name Methanococcus jannaschii
DATE       13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change
10-Oct-1997
ACCESSIONS E64474
REFERENCE   A64300
#authors   Bult, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann,
           R.D.; Sutton, G.G.; Blake, J.A.; FitzGerald, L.M.; Clayton,
           R.A.; Gocayne, J.D.; Kerlavage, A.R.; Dougherty, B.A.;
           Tomb, J.F.; Adams, M.D.; Reich, C.I.; Overbeek, R.;
           Kirkness, E.F.; Weinstock, K.G.; Merrick, J.M.; Glodek, A.;
           Scott, J.L.; Geoghegan, N.S.M.; Weidman, J.F.; Fuhrmann,
           J.L.; Nguyen, D.; Utterback, T.R.; Keiley, J.M.; Peterson,
           J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts,
           K.M.; Hurst, M.A.; Kaine, B.P.; Borodovsky, M.; Klenk,
           H.P.; Fraser, C.M.; Smith, H.O.; Woese, C.R.; Venter, J.C.
#journal   Science (1996) 273:1058-1073
#title     Complete genome sequence of the methanogenic archaeon,
           Methanococcus jannaschii.
#cross-references MUID:96337999
#accession  E64474
#status    preliminary; nucleic acid sequence not shown;
           translation not shown
#molecule_type DNA
#residues  1-391 #label BUL
#cross-references GB:U67580; GB:L77117; NID:g1592044; PID:g1592046;
           TIGR:MJ1398; PID:g1511417
GENETICS    #map_position FOR1358185-1359360
SUMMARY     #length 391 #molecular-weight 45529 #checksum 6984
Query Match   4.7%; Score 98; DB 2; Length 391;
Best Local Similarity 30.0%; Pred. No. 7.17e-01;
Matches 15; Conservative 19; Mismatches 14; Indels 2; Gaps 2;

Db 84 LRFKIKNSYLIITKKNVNTKMAIIFTSINGSRVNGNKTIIWKDPL 133
|| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 176 LKIKVKEITYINRDTKILLETSKTI-YK-LNGVSRDLKKSVLWIKDSL 223

Search completed: Tue Aug 4 10:41:38 1998
Job time : 61 secs.
```

THIS PAGE BLANK (USPTO)



(TW)

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPsrch_pp protein - protein database search, using Smith-Waterman algorithm
Run on: Tue Aug 4 10:41:56 1998; MasPar time 10.08 Seconds
Tabular output not generated. 684.576 Million cell updates/sec

Title: >US-08-848-439-3
Description: (1-275) from US08848439.pep
Perfect Score: 2073
Sequence: 1 SARGLFQGPDFSFKRSNC.....WQKGQREKFRISIRKLQC 275

Scoring table:

PAM 150
Gap 11

Searched: 69111 seqs, 25083644 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: swiss-prot35
1:swiss1

Statistics: Mean 46.355; Variance 72.254; scale 0.642

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	ID	Description	Pred. No.
1	332	16.0	581	FRIZ_DROME	5.63e-52
2	148	7.1	581	NETRIN-2 PRECURSOR (FR	1.08e-10
3	124	6.0	606	NETRIN-1 PRECURSOR.	3.47e-06
4	109	5.3	465	YQ18_CAEEL	1.39e-02
5	101	4.9	1741	DNA-DIRECTED RNA POLYM	2.78e-03
6	100	4.8	500	CYTCHROME P450 IID18	3.99e-02
7	100	4.8	500	CYTCHROME P450 IID4	3.99e-02
8	99	4.8	724	RYBOSOMAL PROTEIN S6 K	5.72e-02
9	98	4.7	1032	BASEPLATE STRUCTURAL P	8.18e-02
10	96	4.6	520	INTERMEDIATE FILAMENT	1.66e-01
11	94	4.5	531	UDP-GLUCURONOSYLTRANSF	3.32e-01
12	92	4.4	1562	HYPOTHETICAL 180.2 KD	6.58e-01
13	90	4.3	216	DNA GYRASE SUBUNIT B	1.29e+00
14	89	4.3	262	VIRUF_SHIDY	1.79e+00
15	90	4.3	263	DIMETHYLADEINOSINE TRAN	1.29e+00
16	90	4.3	561	VACUOLAR ATP SYNTHASE	1.29e+00
17	89	4.3	752	KSA6_CHICK	1.79e+00
18	88	4.2	127	YB99_YEAST	2.48e+00
19	88	4.2	182	HYPOTHETICAL 20.1 KD P	2.48e+00
20	88	4.2	443	ALPHA-AMYLASE ISOZYME	2.48e+00
21	88	4.2	445	ALPHA-AMYLASE ISOZYME	2.48e+00
22	87	4.2	446	HYPOTHETICAL 50.0 KD P	3.43e+00
23	87	4.2	510	ONCONEURAL VENTRAL ANT	3.43e+00

24	87	4.2	580	1	VATA_HORVU	VACUOLAR ATP SYNTHASE	3.43e+00
25	88	4.2	633	1	KS62_MOUSE	RYBOSOMAL PROTEIN S6 K	2.48e+00
26	88	4.2	740	1	KS62_HUMAN	RYBOSOMAL PROTEIN S6 K	2.48e+00
27	88	4.2	848	1	NEM_MOUSE	NEUROFILAMENT TRIPLET	2.48e+00
28	88	4.2	1102	1	YK64_CAEEL	HYPOTHETICAL 124.8 KD	2.48e+00
29	87	4.2	1201	1	COPA_YEAST	COATOMER ALPHA SUBUNIT	3.43e+00
30	87	4.2	1752	1	DESP_HUMAN	DESMOPLAKIN I AND II	3.43e+00
31	87	4.2	2749	1	IP3R_MOUSE	INOSITOL 1,4,5-TRISPHO	3.43e+00
32	87	4.2	2749	1	IP3R_RAT	INOSITOL 1,4,5-TRISPHO	3.43e+00
33	85	4.1	216	1	GYRB_ACIS7	DNA GYRASE SUBUNIT B	6.48e+00
34	86	4.1	327	1	VMP_CAWD	MOVEMENT PROTEIN (CELL	4.72e+00
35	86	4.1	450	1	DCOR_CHICK	ORNITHINE DECARBOXYLASE	4.72e+00
36	86	4.1	461	1	DCOR_BOVIN	ORNITHINE DECARBOXYLASE	4.72e+00
37	85	4.1	495	1	EIBL_ADE02	E1B PROTEIN, LARGE T-A	6.48e+00
38	86	4.1	556	1	FTHS_STRMU	FORMATE--TETRAHYDROFOL	4.72e+00
39	86	4.1	620	1	Y870_METJA	HYPOTHETICAL PROTEIN M	4.72e+00
40	85	4.1	633	1	YHR4_YEAST	HYPOTHETICAL 71.2 KD P	6.48e+00
41	86	4.1	725	1	VACB_MYCGE	VACB PROTEIN HOMOLOG.	4.72e+00
42	85	4.1	798	1	VP16_YEAST	VACUOLAR PROTEIN SORTI	6.48e+00
43	85	4.1	1407	1	CYAA_DICDI	ADENYLATE CYCLASE, AGG	6.48e+00
44	85	4.1	1726	1	MSPI_PLAEP	MEROZOITE SURFACE PROT	6.48e+00
45	85	4.1	2688	1	ZEPI_MOUSE	ZINC FINGER PROTEIN 40	6.48e+00

ALIGNMENTS

RESULT	1	STANDARD;	PRT;	581 AA.
ID	FRIZ_DROME			
AC	P18537;			
DT	01-NOV-1990 (REL. 16, CREATED)			
DT	01-NOV-1990 (REL. 16, LAST SEQUENCE UPDATE)			
DT	01-MAY-1992 (REL. 22, LAST ANNOTATION UPDATE)			
DE	FRIZZLED PROTEIN PRECURSOR.			
GN	FZ			
OS	DROSOPHILA MELANOGASTER (FRUIT FLY).			
OC	EUKARYOTA; METAZOA; ARTHROPODA; INSECTA; DIPTERA.			
RN	[1]			
RP	SEQUENCE FROM N.A. (CLASS I).			
RX	MEDLINE; 89159415.			
RA	VINSON C.R. CONOVER S., ADLER P.N.;			
RL	NATURE 338:263-264(1989).			
RN	[2]			
RP	SEQUENCE FROM N.A. (CLASSES I AND II).			
RX	MEDLINE; 91060073.			
RA	ADLER P.N., VINSON C., PARK W.J., CONOVER S., KLEIN L.;			
RL	GENETICS 126:401-416(1990).			
CC	-1- FUNCTION: REQUIRED TO COORDINATE THE CYTOSKELETONS OF EPIDERMAL			
CC	CELLS TO PRODUCE A PARALLEL ARRAY OF CUTICULAR HAIRS AND BRISTLES.			
CC	FZ IS REQUIRED FOR THE TRANSMISSION OF POLARITY INFORMATION, AND			
CC	TRANSDUCTION.			
CC	-1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.			
CC	-1- ALTERNATIVE PRODUCTS: CLASS I AND CLASS II PROTEINS DIFFER AT THE			
CC	C-TERMINAL FROM AA 406 DUE TO USE OF ALTERNATIVE 3' EXONS.			
DR	EMBL; X54648; G804979; JOINED.			
DR	EMBL; X54649; G804979; JOINED.			
DR	EMBL; X54650; G804979; JOINED.			
DR	EMBL; X54651; G804979; JOINED.			
DR	EMBL; X54648; G804980; JOINED.			
DR	EMBL; X54649; G804980; JOINED.			
DR	EMBL; X54650; G804980; JOINED.			
DR	EMBL; X54652; G804980; JOINED.			
DR	EMBL; X54646; G7981; JOINED.			
DR	EMBL; X54647; G7983; ALT_SEQ.			
DR	PIR; S03540; S03540.			
DR	PIR; S15708; S15708.			
DR	FLYBASE; FBgn0001085; fz.			
KW	TRANSMEMBRANE; SIGNAL; ALTERNATIVE SPLICING; DEVELOPMENTAL PROTEIN.			
FT	SIGNAL 1 26			
FT	POTENTIAL.			
FT	CHAIN 27 581			
FT	FRIZZLED PROTEIN.			
FT	TRANSMEM 248 270			
FT	POTENTIAL.			
FT	TRANSMEM 281 303			
FT	POTENTIAL.			
FT	TRANSMEM 336 368			
FT	POTENTIAL.			
FT	TRANSMEM 381 401			
FT	POTENTIAL.			

FT TRANSMEM 422 439 POTENTIAL.
FT TRANSMEM 472 492 POTENTIAL.
FT TRANSMEM 530 553 POTENTIAL.
FT VARSPLIC 406 415 DILSGVCFVG -> MYLMOFHTIN (IN CLASS II).
FT VARSPLIC 416 581 MISSING (IN CLASS II).
SQ SEQUENCE 581 AA; 64847 MW; 279EBE85 CRC32;

Query Match 16.0%; Score 332; DB 1; Length 581;
Best Local Similarity 37.5%; Pred. No. 5.63e-52;
Matches 45; Conservative 22; Mismatches 47; Indels 6; Gaps 6;

Db 50 HNRCEPI-T-ISICKNPYNMTIMPNGHGTQBEAGLEVHQFAPLYKIGCSDDLQFLC 107
QY 17 RSNCKPAPNLQCHGLEYQNMPLNLLGHETKVEVLEQAGAMPLVYKQCHPDTKKFLC 76
Db 108 SLYVPVC-TILERIPPCRSCEASRY-CEKLMKTYNFPNENLECSKFPVGHGEDICVA 165
QY 77 SLFAPVCLDDLETIOFCHSLCVQVQKRCAPVMSAFGPPWPMLECDRFP-QD-NDLICP 134

RESULT 2
ID NET2_CHICK STANDARD; PRT; 581 AA.
AC Q90923;
DT 01-NOV-1997 (REL. 35, CREATED)
DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE NETRIN-2 PRECURSOR (FRAGMENT).
OS GALLUS GALLUS (CHICKEN).
OC EUKARYOTA; CHORDATA; VERTEBRATA; TETRAPODA; AVES; NEOGNATHAE;
OC GALLIFORMES.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-WHITE LEHORN; TISSUE-EMBRYONIC BRAIN;
RX MEDLINE; 94340732.
RA SERAFINI T., KENNEDY T.E., GALKO M.J., MIRZAYAN C., JESSELL T.M.,
RA TESSIER-LAVIGNE M.;
RL CELL 78:409-424(1994).
CC -!- FUNCTION: NETRINS CONTROL GUIDANCE OF CNS COMMISSURAL AXONS AND
CC PERIPHERAL MOTOR AXONS.
CC -!- SUBCELLULAR LOCATION: EXTRACELLULAR.
CC -!- SIMILARITY: CONTAINS ONE LAMININ N-TERMINAL DOMAIN (DOMAIN VI).
CC -!- SIMILARITY: CONTAINS 3 LAMININ EGF-LIKE DOMAINS.
CC -!- SIMILARITY: CONTAINS ONE C345C DOMAIN.
CC -!- SIMILARITY: TO CAENORHABDITIS ELEGANS UNC-6 PROTEIN.
DR EMBL: L34550; G529421; .
DR PROSITE; PS00022; EGF_1; 2.
DR PROSITE; PS01248; LAMININ_TYPE_EGF; 3.
KW GLYCOPROTEIN; EXTRACELLULAR MATRIX; SIGNAL; LAMININ EGF-LIKE DOMAIN;
KW REPEAT.
FT SIGNAL 1 1
FT NON_TER 1 15
FT CHAIN <1 15
FT DOMAIN 16 581
FT DOMAIN 16 261
FT DOMAIN 262 430
FT DOMAIN 262 317
FT DOMAIN 318 380
FT DOMAIN 381 430
FT DOMAIN 431 581
FT SITE 507 509
FT SITE 507 124
FT DISULFID 91 124
FT DISULFID 262 271
FT DISULFID 264 281
FT DISULFID 283 292
FT DISULFID 295 315
FT DISULFID 318 327
FT DISULFID 320 345
FT DISULFID 348 357
FT DISULFID 360 378
FT DISULFID 381 393
FT DISULFID 383 400
FT DISULFID 402 411
FT DISULFID 414 428

FT DISULFID 453 521 BY SIMILARITY.
FT DISULFID 458 578 BY SIMILARITY.
FT CARBOHYD 88 88 POTENTIAL.
FT CARBOHYD 103 103 POTENTIAL.
FT CARBOHYD 394 394 POTENTIAL.
FT CARBOHYD 540 540 POTENTIAL.
SQ SEQUENCE 581 AA; 65106 MW; 5B6D2272 CRC32;
Query Match 7.1%; Score 148; DB 1; Length 581;
Best Local Similarity 27.0%; Pred. No. 1.06e-10;
Matches 30; Conservative 28; Mismatches 46; Indels 7; Gaps 6;

Db 431 IPAINPTSLVSTPE-APADCDYSCKPAKGYKINMKKYCKDYVQVNVILEMETVANWAK 489
QY 133 IPLASSDHLPLATEAPKVCFA-CKNKNDNDNDIMETLCKNDFALAIKVKYKEIYINRDK 191
Db 490 FTINILS-VYKORDRVRKR-GDNFLMIHLKDLSCPKI-QISKYLYMG 536
QY 192 IILETRSKTIYKLVGSRDLAKSVLWLK-DSLQCTCEEMNDINAPLYMG 241

RESULT 3
ID NET1_CHICK STANDARD; PRT; 606 AA.
AC Q90922;
DT 01-NOV-1997 (REL. 35, CREATED)
DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE NETRIN-1 PRECURSOR.
OS GALLUS GALLUS (CHICKEN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; AVES; NEOGNATHAE;
OC GALLIFORMES.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-WHITE LEHORN; TISSUE-EMBRYONIC BRAIN;
RX MEDLINE; 94340732.
RA SERAFINI T., KENNEDY T.E., GALKO M.J., MIRZAYAN C., JESSELL T.M.,
RA TESSIER-LAVIGNE M.;
RL CELL 78:409-424(1994).
CC -!- FUNCTION: NETRINS CONTROL GUIDANCE OF CNS COMMISSURAL AXONS AND
CC PERIPHERAL MOTOR AXONS.
CC -!- SUBCELLULAR LOCATION: EXTRACELLULAR.
CC -!- SIMILARITY: CONTAINS ONE LAMININ N-TERMINAL DOMAIN (DOMAIN VI).
CC -!- SIMILARITY: CONTAINS 3 LAMININ EGF-LIKE DOMAINS.
CC -!- SIMILARITY: CONTAINS ONE C345C DOMAIN.
CC -!- SIMILARITY: TO CAENORHABDITIS ELEGANS UNC-6 PROTEIN.
DR EMBL: L34549; G529419; .
DR PROSITE; PS00022; EGF_1; 2.
DR PROSITE; PS01248; LAMININ_TYPE_EGF; 3.
KW GLYCOPROTEIN; EXTRACELLULAR MATRIX; SIGNAL; LAMININ EGF-LIKE DOMAIN;
KW REPEAT.
FT SIGNAL 1 25
FT CHAIN 26 606
FT DOMAIN 26 286
FT DOMAIN 287 455
FT DOMAIN 287 342
FT DOMAIN 343 405
FT DOMAIN 406 455
FT DOMAIN 456 606
FT SITE 532 534
FT SITE 532 121
FT DISULFID 121 154
FT DISULFID 287 296
FT DISULFID 289 306
FT DISULFID 308 317
FT DISULFID 320 340
FT DISULFID 343 352
FT DISULFID 345 370
FT DISULFID 373 382
FT DISULFID 385 403
FT DISULFID 406 418
FT DISULFID 408 425
FT DISULFID 427 436
FT DISULFID 439 453

	CC	OF DNA INTO RNA USING THE FOUR RIBONUCLEOSIDE TRIPHOSPHATES AS SUBSTRATES.
-!	CC	CATALYTIC ACTIVITY: N NUCLEOSIDE TRIPHOSPHATE = N PYROPHOSPHATE + RNA(N).
-!	CC	SUBUNIT: RNA POLYMERASE III CONSISTS OF ABOUT 15 DIFFERENT SUBUNITS. THIS SUBUNIT IS THE LARGEST COMPONENT OF RNA POLYMERASE III.
-!	CC	SUBCELLULAR LOCATION: NUCLEAR.
-!	CC	THREE DISTINCT ZINC-CONTAINING RNA POLYMERASES ARE FOUND IN EUKARYOTIC NUCLEI: POLYMERASE I FOR THE RIBOSOMAL RNA PRECURSOR, POLYMERASE II FOR THE MRNA PRECURSOR, AND POLYMERASE III FOR 5S AND TRNA GENES.
EMBL;	CC	X60325; G9364; -.
DNA-DIRECTED	DR	RNA POLYMERASE; TRANSCRIPTION; ZINC; ZINC-FINGER; NUCLEAR PROTEIN.
ZNFING	KW	79 C2H2-TYPE (POTENTIAL).
SEQUENCE	FF	1741 AA; 193928 MW; DCDCD942 CRC32;
Query Match	SQ	4.9%; Score 101; DB 1; Length 1741;
Best Local Similarity		27.4%; Pred. No. 2.78e-02;
Matches		20; Conservative 25; Mismatches 23; Indels 5; Gaps 5;
Db	168	TICKSQGFRV-LREMHVEEARATKLADVIKKDKIEKIQLDQR-LRNHHVVDPQAOLHI 225 : : : : : : : : : : : : : : : : :
QY	168	TICKNDFAUKIVKVTETIYNRDTKILET-KS-KTYILNGVSEDLKKSVLWLKDLOC 225 : : : : : : : : : : : : : : : : :
Db	226	L-QKVPECDPYL 237 : : : :
QY	226	TCEENNDINAPYL 238 : : : :
RESULT	6	STANDARD; PRT: 500 AA.
ID	CPDI	RAT Q64680;
AC	Q64680;	
DT	01-NOV-1997	(REL. 35, CREATED)
DT	01-NOV-1997	(REL. 35, LAST SEQUENCE UPDATE)
DT	01-NOV-1997	(REL. 35, LAST ANNOTATION UPDATE)
DE	CYCITROME P450 IID18	(BC 1.14.14.1) (P450 2D-29/2D-35). CYPD218 OR CYP2D-18.
OS	RATTUS NORVEGICUS	(RAT).
OC	EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA; EUTHERIA; RODENTIA. [1]	
RN	SEQUENCE FROM N.A.	
RP	MEDLINE:	95251650.
RX	KAWASHIMA H.; STROBEL H.W.	
RL	BIOCHEM. BIOPHYS. RES. COMMUN.	209:535-540(1995).
CC	-!	FUNCTION: CYTOCHROMES P450 ARE A GROUP OF HEME-THIOLATE MONOOXYGENASES. IN LIVER MICROSOMES, THIS ENZYME IS INVOLVED IN AN NAUPH-DEPENDENT ELECTRON TRANSPORT PATHWAY. IT OXIDIZES A VARIETY OF STRUCTURALLY UNRELATED COMPOUNDS, INCLUDING STEROIDS, FATTY ACIDS, AND XENOBIOTICS.
CC	-!	CATALYTIC ACTIVITY: RH + REDUCED FLAVOPROTEIN + O(2) - ROH + OXIDIZED FLAVOPROTEIN + H(2)O.
CC	-!	SUBCELLULAR LOCATION: MEMBRANE-BOUND.
CC	-!	TISSUE SPECIFICITY: BRAIN.
CC	-!	SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
DR	EMBL;	U48220; GI200518; -.
DR	EMBL;	U48219; GI200516; -.
DR	PROSITE;	PS00086; CYTOCHROME_P450; 1.
KW	OXIDOREDUCTASE; MONOOXYGENASE; ELECTRON TRANSPORT; MEMBRANE; HEME; MICROsome.	
FT	BINDING	446 446 HEME (BY SIMILARITY).
SQ	SEQUENCE	500 AA; 56683 MW; F0194995 CRC32;
Query Match		4.8%; Score 100; DB 1; Length 500;
Best Local Similarity		26.7%; Pred. No. 3.99e-02;
Matches		16; Conservative 22; Mismatches 18; Indels 4; Gaps 3;
Db	173	FPFSS-PNTLLKAVCNVTASLIIFACR-FEYNDPRIRLL--DLLKDTLEECSGFPLMLLN 228 : : : : : : : : : : : : : : : : :
QY	6	FLFGPDFSKRSCKNPANQLCGHEGYQMRLPNLLGHETMKREVLFQAQAIPLVMK 65 : : : : : : : : : : : : : : : : :

```

RL MOL. CELL. BIOL. 9:3850-3859(1989).
CC -1- FUNCTION: PHOSPHORYLATES A WIDE RANGE OF SUBSTRATES INCLUDING
CC -1- RIBOSOMAL PROTEIN S6.
CC -1- TISSUE SPECIFICITY: INTESTINE, THYMUS, AND LUNG.
CC -1- SIMILARITY: THE N-TERMINUS SHOWS HIGH SEQUENCE SIMILARITY TO THE
CC REGIONS OF PROTEIN KINASE C, THE CATALYTIC SUBUNIT OF CAMP-
CC DEPENDENT PROTEIN KINASE, AND CGMP-DEPENDENT PROTEIN KINASE. THE
CC REMINDER OF S6 KINASE II SHOWS HIGH SEQUENCE SIMILARITY TO THE
CC CATALYTIC SUBUNIT OF PHOSPHORYLASE B KINASE.
CC EMBL; M28489; G556322; -.
DR PIR; B32571; B32571.
DR HSP; P05132; ICTP.
DR MGD; MG1:104558; RPS6KAL.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 2.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 2.
DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 2.
DR TRANSFERASE; SERINE/THREONINE-PROTEIN KINASE; ATP-BINDING;
KW REPEAT; MULTIGENE FAMILY.
FT FT DOMAIN 62 310 PROTEIN KINASE 1.
FT FT DOMAIN 407 664 PROTEIN KINASE 2.
FT FT NP_BIND 68 76 ATP (BY SIMILARITY).
FT BINDING 94 94 ATP (BY SIMILARITY).
FT ACT_SITE 187 187 BY SIMILARITY.
FT FT NP_BIND 413 421 ATP (BY SIMILARITY).
FT BINDING 436 436 ATP (BY SIMILARITY).
FT ACT_SITE 524 524 BY SIMILARITY.
SQ SEQUENCE 724 AA; 81594 MW; F961A261 CRC32;

Query Match 4.8%; Score 99; DB 1; Length 724;
Best Local Similarity 35.2%; Pred.No. 5.72e-02;
Matches 19; Conservative 13; Mismatches 18; Indels 4; Gaps 4;

Db 489 KI-LROKFFSERERASFLVHTFISKTVELYHSQGVVHRLKPSNITLYVDESGNPEC 541
|| :::: ::::: ||||: || ||||| ::::: ||: ||: ||
QY 177 KIKVKEITYINRDTKILETKSII-Y-KINGVSRDLKKS-VIWLKDSLQCTC 227

RESULT 9
ID VQ07_BPT4 STANDARD; PRT; 1032 AA.
AC P19061;
AD 01-NOV-1990 (REL. 16, CREATED)
DT 01-NOV-1990 (REL. 16, LAST SEQUENCE UPDATE)
DT 01-NOV-1990 (REL. 16, LAST SEQUENCE UPDATE)
DE BASEPLATE STRUCTURAL PROTEIN GP7.
GN 7.
OS BACTERIOPHAGE T4.
OC VIRIDAE; DS-DNA NONENVELOPED VIRUSES; MYOVIRIDAE.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=D;
RX MEDLINE; 90384864.
RA EFIMOV V.P., PRILIPOV A.G., MESYANZHINOV V.V.;
RL NUCLEIC ACIDS RES. 18:5313-5313(1990).
CC -1- FUNCTION: STRUCTURAL COMPONENT OF THE BASEPLATE.
DR EMBL; X15907; G15323; -.
DR PIR; JQ0657; G7BPT4.
KW STRUCTURAL PROTEIN.
SQ SEQUENCE 1032 AA; 119214 MW; 39F5B2D1 CRC32;

Query Match 4.7%; Score 98; DB 1; Length 1032;
Best Local Similarity 39.0%; Pred.No. 8.18e-02;
Matches 16; Conservative 8; Mismatches 15; Indels 2; Gaps 2;

Db 368 MDSICDKVFALIGEVETLNAPRTSKII-DSADKGIIYVLN 407
||::: ||| | : : ||| :: |||
QY 166 METLCKNDFAIKI-KVKEITYINRDTKIILETKSKTIYKLN 205

RESULT 10
ID ION3_CARAU STANDARD; PRT; 520 AA.
AC F18520;
AD 01-NOV-1990 (REL. 16, CREATED)
DT 01-NOV-1990 (REL. 16, LAST SEQUENCE UPDATE)

```

```

DR EMBL; N34007; G207579; ALT_TERM.
DR POSITE; P500375; UDPGT; 1.
KW TRANSFERASE; GLYCOSYLTRANSFERASE; GLYCOPROTEIN; TRANSMEMBRANE; SIGNAL;
KW MULTIGENE FAMILY; MICROSOME; ALTERNATIVE SPLICING.
FT SIGNAL 1 25 POTENTIAL.
FT CHAIN 26 531 UDP-GLUCURONOSYLTFRANSFERASE 1-3.
FT TRANSMEM 489 505 POTENTIAL.
FT CARBOHYD 116 116 POTENTIAL.
FT CARBOHYD 139 139 POTENTIAL.
FT CARBOHYD 293 293 POTENTIAL.
FT CARBOHYD 431 431 POTENTIAL.
SQ SEQUENCE 531 AA; 60140 MW; C3B2A846 CRC32;

Query Match 4.5%; Score 94; DB 1; Length 531;
Best Local Similarity 35.7%; Pred.No. 3.32e-01;
Matches 15; Conservative 15; Mismatches 7; Indels 5; Gaps

Db 214 PVP-WMYTLC-VNYGSLARLASDLLQREVSVVVEILRHASVWL 253
| | : | | : | | : | | : | | : | | : | | : | | : | | : | | : |
Qy 22 PIPANQLCHGIEYQN-RLP-NLLGHE-TMKEVLEQAGAWI 60
| | : | | : | | : | | : | | : | | : | | : | | : | | : | | : |

RESULT 12
ID YN81-YEAST STANDARD; PRT: 1562 AA.
AC Q04781; Q04029;
DT 01-NOV-1997 (REL. 35, CREATED)
DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
DE 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE HYPOTHETICAL 180.2 KD PROTEIN IN FAA4-HOR7 INTERGENIC REGION.
GN YNR247C OR YMR9408.09C OR YMR920.01C.
OS SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).
OC EUKARYOTA; FUNGI; ASCOMYCOTINA; HEMIASCOMYCETES.
[1]
RN RP SEQUENCE OF 1-956 FROM N.A.
RC STRAIN=5288C / AB972;
RA HUNT S., BOWMAN S.; BARRELL B.G.; RAJANDREAM M.A.; WALSH S.V.;
RL SUBMITTED (MAR-1995) TO EMBL/GENBANK/DDBJ DATA BANKS.
[2]
RN RP SEQUENCE OF 950-1562 FROM N.A.
RC STRAIN=5288C / AB972;
RA GENTLES S., BOWMAN S.; BARRELL B.G.; RAJANDREAM M.A.; WALSH S.V.;
RL SUBMITTED (MAR-1995) TO EMBL/GENBANK/DDBJ DATA BANKS.
DR EMBL; Z48639; G732925; -
DR EMBL; Z48756; G736313; -
KW HYPOTHETICAL PROTEIN.
SQ SEQUENCE 1562 AA; 180185 MW; CEE45672 CRC32;

Query Match 4.4%; Score 92; DB 1; Length 1562;
Best Local Similarity 38.8%; Pred.No. 6.58e-01;
Matches 19; Conservative 9; Mismatches 21; Indels 0; Gaps

Db 1397 KNEFDINSKMRLTSNDALTIKLNNITNEVKASYLDQQKLEISFKL 1445
| | : | | : | | : | | : | | : | | : | | : | | : | | : | | : |
Qy 156 KNKDDNDIMETLCKNDFAIKYKEITYINRDTKIILETKSKTYIKL 204
| | : | | : | | : | | : | | : | | : | | : | | : | | : | | : |

RESULT 13
ID GYRB-ACIS3 STANDARD; PRT: 216 AA.
AC Q04273; Q60168;
DT 01-NOV-1997 (REL. 35, CREATED)
DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
DE 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE DNA GYRASE SUBUNIT B (EC 5.99.1.3) (FRAGMENTS).
GN GYRB.
OS ACINETOBACTER SP. (STRAIN ATCC 17924 / CIP 70.12).
OC PROKARYOTA; GRACILICUTES; SCOTOBACTERIA; AEROBIC RODS AND COCCI;
OC NEISSERIACEAE.
[1]
RN RP SEQUENCE FROM N.A.
RX MEDLINE; 9708899.
RA YAMAMOTO S., HARAYAMA S.;
RL INT. J. SYST. BACTERIOL. 46:506-511(1996).
CC -!- FUNCTION: DNA GYRASE NEGATIVELY SUPERCOILS CLOSED CIRCULAR DOUBLE

```

CC STRANDED DNA IN AN ATP-DEPENDANT MANNER AND ALSO CATALYZES THE
 CC INTERCONVERSION OF OTHER TOPOLOGICAL ISOMERS OF DOUBLE-STRANDED
 CC DNA RINGS, INCLUDING CATENANES AND KNOTTED RINGS.
 CC -!- CATALYTIC ACTIVITY: ATP-DEPENDENT BREAKAGE, PASSAGE AND REJOINING
 CC OF DOUBLE-STRANDED DNA.
 CC -!- SUBUNIT: MADE UP OF TWO CHAINS. THE A CHAIN IS RESPONSIBLE FOR DNA
 CC BREAKAGE AND REJOINING; THE B CHAIN CATALYZES ATP HYDROLYSIS. THE
 CC ENZYME FORMS AN A2B2 TETRAMER.
 DR EMBL: D73436; G1322131; -.
 DR EMBL: D73421; G1322101; -.
 DR PROSITE: PS00177; TOPOISOMERASE_II; 1.
 KW TOPOISOMERASE: ISOMERASE; ATP-BINDING; ANTIBIOTIC RESISTANCE.
 FT NON_TER 1 1
 FT NON_CONS 116 117
 FT NON_TER 216 216
 SQ SEQUENCE 216 AA; 23775 MW; E572BEE0 CRC32;

Query Match 4.3%; Score 90; DB 1; Length 216;
 Best Local Similarity 26.3%; Pred. No. 1.29e+00;
 Matches 20; Conservative 23; Mismatches 26; Indels 7; Gaps 7;

Db 79 LARRURELSFLNAGVRIYLRDERVALEHFDVGLSEKSL-DIAGLPGLA-DCOEKD 136
 QY 176 LKIKVKEITYINRDKIIL-EIK-S-KTIYKLN-GVSEDLKKSVLWLKDSLOCTCEMN 231
 Db 137 PALSLEYLVEGDSAGG 152
 QY 232 -DINAPYLVMGOKGG 246

RESULT 14
 ID VIRF_SHIDY STANDARD; PRT; 262 AA.
 AC Q04248;
 DT 01-JUN-1994 (REL. 29, CREATED)
 DT 01-JUN-1994 (REL. 29, LAST SEQUENCE UPDATE)
 DT 01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)
 DE VIRULENCE REGULON TRANSCRIPTIONAL ACTIVATOR VIRF.
 GN VIRF.
 OS SHIGELLA DYSENTERIAE, SHIGELLA FLEXNERI, AND SHIGELLA SONNEI.
 OG PLASMID 210 KB INVASION.
 OC PROKARYOTA; GRACILICUTES; SCOTOBACTERIA; FACULTATIVELY ANAEROBIC RODS;
 OC ENTEROBACTERIACEAE.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC SPECIES=S.DYSENTERIAE; STRAIN=CG097;
 RA YAO R., REDDY L.V., PALCHAUDHURI S.;
 RL SUBMITTED (MAR-1991) TO EMBL/GENBANK/DBJ DATA BANKS.
 RP [2]
 RP SEQUENCE FROM N.A.
 RC SPECIES=S.FLEXNERI; STRAIN=2A;
 RX MEDLINE; 87032409.
 RA SAKAI T., SASAKAWA C., MAKINO S., YOSHIKAWA M.;
 RL INFECT. IMMUN. 54:395-402(1986).
 RP [3]
 RP SEQUENCE FROM N.A.
 RC SPECIES=S.SONNEI;
 RX MEDLINE; 89212891.
 RA KATO J.I., ITO K.I., NAKAMURA A., WATANABE H.;
 RL INFECT. IMMUN. 57:1391-1398(1989).
 RP [4]
 RP SIMILARITY TO ARAC FAMILY.
 RX MEDLINE; 92326642.
 RA DORMAN C.J.;
 RL MOL. MICROBIOL. 6:1575-1575(1992).
 CC -!- FUNCTION: TRANSCRIPTIONAL ACTIVATOR OF THE VIRB GENE WHICH IS
 CC ITSELF AN ACTIVATOR OF THE IPAABC VIRULENCE REGULON.
 CC -!- SIMILARITY: BELONGS TO THE ARAC/XLYS FAMILY OF TRANSCRIPTIONAL
 CC REGULATORS.
 DR EMBL: X58464; G46954; -.
 DR EMBL: M29172; G152801; -.
 DR EMBL: X16661; G47067; -.
 DR PIR: S14646; S14646.
 DR PIR: A47605; A47605.

DR PIR: A60105; A60105.
 DR PROSITE: PS00041; HTH_ARAC_FAMILY_1; 1.
 DR PROSITE: PS01124; HTH_ARAC_FAMILY_2; 1.
 KW PLASMOD: VIRULENCE; TRANSCRIPTION REGULATION; ACTIVATOR; DNA-BINDING.
 FT DNA_BIND 177
 SQ SEQUENCE 262 AA; 30549 MW; 11339B38 CRC32;

Query Match 4.3%; Score 89; DB 1; Length 262;
 Best Local Similarity 25.5%; Pred. No. 1.79e+00;
 Matches 13; Conservative 21; Mismatches 15; Indels 2; Gaps 2;

Db 37 TLTIDEQIAFIERNQINQVSIKKSINPFETISLDNRLLLSIIRIMEPI 87
 QY 175 ALKIKVKEITYINRDKIILETK-SKTIYKLVGS-ERDLKKSVLWLKDSL 223

RESULT 15
 ID KSGA_MYCPN STANDARD; PRT; 263 AA.
 AC P75113;
 DT 01-NOV-1997 (REL. 35, CREATED)
 DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
 DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
 DE DIMETHYLADENOSINE TRANSFERASE (EC 2.1.1.-) (S-ADENOSYLMETHIONINE-6-N',
 DE N'-ADENOSYL(RRNA) DIMETHYLTRANSFERASE) (16S RRNA DIMETHYLASE) (HIGH
 DE LEVEL KASUGAMYCIN RESISTANCE PROTEIN KSGA) (KASUGAMYCIN
 DE DIMETHYLTRANSFERASE).
 GN KSGA.
 OS MYCOPLASMA PNEUMONIAE.
 OC PROKARYOTA; TENERICUTES; MOLLICUTES; MYCOPLASMA; MYCOPLASMATALES;
 OC MYCOPLASMATACEAE.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 29342 / M129;
 RX MEDLINE; 97105885.
 RA HIMMELREICH R., HILBERT H., PLAGENS H., PIRKL E., LI B.-C.,
 RA HERRMANN R.;
 RL NUCLEIC ACIDS RES. 24:4420-4449(1996).
 CC -!- FUNCTION: SPECIFICALLY DIMETHYLATES TWO ADJACENT ADENOSINES IN THE
 CC LOOP OF A CONSERVED HAIRPIN NEAR THE 3'-END OF 16S RRNA IN THE 30S
 CC PARTICLE. ITS INACTIVATION LEADS TO KASUGAMYCIN RESISTANCE
 CC (BY SIMILARITY).
 CC -!- SIMILARITY: STRONG, TO OTHER BACTERIAL KSGA.
 CC -!- SIMILARITY: BELONGS TO THE RRNA ADENINE N-6-METHYLTRANSFERASE
 CC FAMILY.
 DR EMBL: AE000017; G1673824; -.
 DR PROSITE: PS01131; RRNA_A_DIMETH; 1.
 KW MRNA PROCESSING: TRANSFERASE; METHYLTRANSFERASE;
 KW ANTIBIOTIC RESISTANCE.
 SQ SEQUENCE 263 AA; 29794 MW; 357AA00D CRC32;

Query Match 4.3%; Score 90; DB 1; Length 263;
 Best Local Similarity 37.0%; Pred. No. 1.29e+00;
 Matches 20; Conservative 11; Mismatches 20; Indels 3; Gaps 3;

Db 52 KLQLPYHGIEL-DKRLAEVLLVNEILTEQLTIGDALKONLDQYFDDTPIPLCG 104
 QY 26 NLQL-CHGIEYQNMRLPN-LLGHETMKEVLEQAGAWIPLVMKQCHPDTRKFLCS 77

Search completed: Tue Aug 4 10:42:21 1998
 Job time : 25 secs.


```
RESULT 2
ID O35297 PRELIMINARY; PRT; 295 AA.
AC O35297;
DT 01-JAN-1998 (TREMBLREL. 05, CREATED)
DT 01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)
DT 01-JAN-1998 (TREMBLREL. 05, LAST ANNOTATION UPDATE)
DE SECRETED APOPTOSIS RELATED PROTEIN 1.
GN SARPL
OS MUS MUSCULUS (MOUSE).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; RODENTIA.
RN [1]
RP SEQUENCE FROM N.A.
RA MELKONYAN H., CHANG W.C., SHAPIRO J.P., MAHADEVAPPA M.,
RA FITZPATRICK P.A., KIEFER M.C., TOMEI D.L., UMANSKY S.R.;
RL SUBMITTED (SEP-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL; AF017989; G2415421;
SQ SEQUENCE 295 AA; 33483 MW; 0E4C8A8E CRC32;

Query Match 99.6%; Score 2065; DB 10; Length 295;
Best Local Similarity 98.9%; Pred. No. 0.00e+00;
Matches 272; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Db 21 SARGLFQGPDPFSYKRSNCKPIPANLQCHGIEYQNMRLPNLLGHETMKEVLEQAGAWI 80
QY 1 SARGLFQGPDPFSYKRSNCKPIPANLQCHGIEYQNMRLPNLLGHETMKEVLEQAGAWI 60

Db 81 PLVMKQCHPDTKKFLCSLFAVPCLDLDLDTIQPCHSLCVQVKDRCAPVMSAF--PWPDM 140
QY 61 PLVMKQCHPDTKKFLCSLFAVPCLDLDLDTIQPCHSLCVQVKDRCAPVMSAF--PWPDM 120

Db 141 ECDRFPQDNDLCIPLASSDHLPLATEAPKVCACKTKNEDNDIMETLCKNDFALKIKV 200
QY 121 ECDRFPQDNDLCIPLASSDHLPLATEAPKVCACKTKNEDNDIMETLCKNDFALKIKV 180

Db 201 KEITYINRDTKIILETSKTIYKLVGSRDLKSVLWLDKSLQCTCEEMNDINAPYLVM 260
QY 181 KEITYINRDTKIILETSKTIYKLVGSRDLKSVLWLDKSLQCTCEEMNDINAPYLVM 240

Db 261 GQGGELVITSVKRWQKQREFKRISRSIRKLOC 295
QY 241 GQGGELVITSVKRWQKQREFKRISRSIRKLOC 275

Db 81 PLVMKQCHPDTKKFLCSLFAVPCLDLDLDTIQPCHSLCVQVKDRCAPVMSAF--PWPDM 140
QY 61 PLVMKQCHPDTKKFLCSLFAVPCLDLDLDTIQPCHSLCVQVKDRCAPVMSAF--PWPDM 120

Db 141 ECDRFPQDNDLCIPLASSDHLPLATEAPKVCACKTKNEDNDIMETLCKNDFALKIKV 200
QY 121 ECDRFPQDNDLCIPLASSDHLPLATEAPKVCACKTKNEDNDIMETLCKNDFALKIKV 180

Db 201 KEITYINRDTKIILETSKTIYKLVGSRDLKSVLWLDKSLQCTCEEMNDINAPYLVM 260
QY 181 KEITYINRDTKIILETSKTIYKLVGSRDLKSVLWLDKSLQCTCEEMNDINAPYLVM 240

Db 261 GQGGELVITSVKRWQKQREFKRISRSIRKLOC 295
QY 241 GQGGELVITSVKRWQKQREFKRISRSIRKLOC 275

RESULT 3
ID P97299 PRELIMINARY; PRT; 295 AA.
AC P97299;
DT 01-MAY-1997 (TREMBLREL. 03, CREATED)
DT 01-MAY-1997 (TREMBLREL. 03, LAST SEQUENCE UPDATE)
DT 01-JAN-1998 (TREMBLREL. 05, LAST ANNOTATION UPDATE)
DE STROMAL CELL DERIVED FACTOR 5 (SDF5).
GN SDF5.
OS MUS MUSCULUS (MOUSE).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; RODENTIA.
RN [1]
RP SEQUENCE FROM N.A.
RA SHIROZU M., TADA H., TASHIRO K., NAKAMURA T., LOPEZ N.D.,
RA NAZAREA M., HAMADA T., SATO T., NAKANO T., HONJO T.;
RL GENOMICS 37:273-280(1996).
DR EMBL; D50462; G1747302;
DR MGD; MGI:108078; SDF5.
SQ SEQUENCE 295 AA; 33501 MW; 329888C8 CRC32;

Query Match 99.5%; Score 2062; DB 10; Length 295;
Best Local Similarity 98.9%; Pred. No. 0.00e+00;
Matches 272; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Db 21 SARGLFQGPDPFSYKRSNCKPIPANLQCHGIEYQNMRLPNLLGHETMKEVLEQAGAWI 80
QY 1 SARGLFQGPDPFSYKRSNCKPIPANLQCHGIEYQNMRLPNLLGHETMKEVLEQAGAWI 60

Db 81 PLVMKQCHPDTKKFLCSLFAVPCLDLDLDTIQPCHSLCVQVKDRCAPVMSAF--PWPDM 140
```

```
QY 61 PLVMKQCHPDTKKFLCSLFAVPCLDLDLDTIQPCHSLCVQVKDRCAPVMSAF--PWPDM 120
Db 141 ECDRFPQDNDLCIPLASSDHLPLATEAPKVCACKTKNEDNDIMETLCKNDFALKIKV 200
QY 121 ECDRFPQDNDLCIPLASSDHLPLATEAPKVCACKTKNEDNDIMETLCKNDFALKIKV 180
Db 201 KEITYINRDTKIILETSKTIYKLVGSRDLKSVLWLDKSLQCTCEEMNDINAPYLVM 260
QY 181 KEITYINRDTKIILETSKTIYKLVGSRDLKSVLWLDKSLQCTCEEMNDINAPYLVM 240
Db 261 GQGGELVITSVKRWQKQREFKRISRSIRKLOC 295
QY 241 GQGGELVITSVKRWQKQREFKRISRSIRKLOC 275

RESULT 4
ID O14778 PRELIMINARY; PRT; 206 AA.
AC O14778;
DT 01-JAN-1998 (TREMBLREL. 05, CREATED)
DT 01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)
DT 01-JAN-1998 (TREMBLREL. 05, LAST ANNOTATION UPDATE)
DE SECRETED APOPTOSIS RELATED PROTEIN 1 (FRAGMENT).
GN SARPL.
OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; PRIMATES.
RN [1]
RP SEQUENCE FROM N.A.
RA MELKONYAN H., CHANG W.C., SHAPIRO J.P., MAHADEVAPPA M., FITZPATRICK P.A.,
RA KIEFER M.C., TOMEI D.L., UMANSKY S.R.;
RL SUBMITTED (SEP-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL; AF017986; G2415415;
FT NON_TER 206
SQ SEQUENCE 206 AA; 23215 MW; F465B68B CRC32;

Query Match 66.4%; Score 1377; DB 2; Length 206;
Best Local Similarity 98.4%; Pred. No. 0.00e+00;
Matches 185; Conservative 0; Mismatches 1; Indels 2; Gaps 1;

Db 21 SARGLFQGPDPFSYKRSNCKPIPANLQCHGIEYQNMRLPNLLGHETMKEVLEQAGAWI 80
QY 1 SARGLFQGPDPFSYKRSNCKPIPANLQCHGIEYQNMRLPNLLGHETMKEVLEQAGAWI 60

Db 81 PLVMKQCHPDTKKFLCSLFAVPCLDLDLDTIQPCHSLCVQVKDRCAPVMSAF--PWPDM 138
QY 61 PLVMKQCHPDTKKFLCSLFAVPCLDLDLDTIQPCHSLCVQVKDRCAPVMSAF--PWPDM 120

Db 139 ECDRFPQDNDLCIPLASSDHLPLATEAPKVCACKTKNEDNDIMETLCKNDFALKIKV 198
QY 121 ECDRFPQDNDLCIPLASSDHLPLATEAPKVCACKTKNEDNDIMETLCKNDFALKIKV 180

Db 199 KEITYINR 206
QY 181 KEITYINR 188

RESULT 5
ID O14779 PRELIMINARY; PRT; 314 AA.
AC O14779;
DT 01-JAN-1998 (TREMBLREL. 05, CREATED)
DT 01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)
DT 01-JAN-1998 (TREMBLREL. 05, LAST ANNOTATION UPDATE)
DE SECRETED APOPTOSIS RELATED PROTEIN 2.
GN SARP2.
OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; PRIMATES.
RN [1]
RP SEQUENCE FROM N.A.
RA TISSUE-HEART;
RA MELKONYAN H., PROCHAZKA V., CHANG W.C., UMANSKY S.R.;
RL SUBMITTED (SEP-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
```


ID	019116	PRELIMINARY;	PRT;	308 AA.
AC	019116;			
DT	01-JAN-1998	(TREMBREL. 05, CREATED)		
DT	01-JAN-1998	(TREMBREL. 05, LAST SEQUENCE UPDATE)		
DT	01-JAN-1998	(TREMBREL. 05, LAST ANNOTATION UPDATE)		
DE	FRZA PRECURSOR.			
DE	BOS TAURUS (BOVINE).			
OC	EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;			
OC	EUTHERIA; ARTIODACTYLIA.			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=HEART;			
RA	DUPLA C., D'AMORE P.A.;			
RL	SUBMITTED (AUG-1997) TO EMBL/GENBANK/DBJ DATA BANKS.			
DR	EMBL; U85945; G2337937; -			
KW	SIGNAL.			
FT	5 22 POTENTIAL.			
SEQ	SEQUENCE 308 AA; 34763 MW; 9B74A92F CRC32;			
	Query Match 43.5%; Score 902; DB 4; Length 308;			
	Best Local Similarity 42.2%; Pred. No. 5.17e-190;			
	Matches 111; Conservative 66; Mismatches 79; Indels 7; Gaps 5;			
Db	45 FYTKPQCVDIPADRLCHLNVGKRMVLPNLLHEHETMAEVKQOASSWPLLNKNCHIGTQ 104			
QY	13 FSYKRSCKPIPANLQCHGIEYQNRLPNLLGHETMKVELEQAGAWIPLVMKQCHPDK 72			
Db	105 VFICSLFAPVCLD---RPIYPCRWLCEAYRDSCEPYMQFFGFYWPMLKCDKFF-EGDYC 160			
QY	73 KFLCSLFAPVCLDDLBETOPCHSLCVQVKRCAPYSAFGFPWDMLECDREFQDNCLC 132			
Db	161 IAMTPPNATEASKPQGTTCVPCPNDELKSEA-IIIEHLCASEFALRMKIKVEVKENGDKKI 219			
QY	133 IPLASDHLLPATEAPKVEACEACKNKNDNDNDIMETLCKNDFAALKIKVKEITYINRDKI 192			
Db	220 V-PKKKKPL-KLGPICKKELKLVLYLKNAGDCPCQHLONLSHHFLIMGKVKVQVLLTA 277			
QY	193 ILTKSKTIYKLVGVSERDLKKSVMKWSLQTCCEMDINDINAPYLVMGQKGGELVITS 252			
Db	278 IHKWDKKNKEFKTFMKMKNHCEC 300			
QY	253 VKRWQKQREFFKISRIRKLQC 275			
RESULT	8			
ID	014780	PRELIMINARY;	PRT;	317 AA.
AC	014780;			
DT	01-JAN-1998	(TREMBREL. 05, CREATED)		
DT	01-JAN-1998	(TREMBREL. 05, LAST SEQUENCE UPDATE)		
DT	01-JAN-1998	(TREMBREL. 05, LAST ANNOTATION UPDATE)		
DE	SICATED APOPTOSIS RELATED PROTEIN 3.			
GN	SARP3.			
OS	HOMO SAPIENS (HUMAN).			
OC	EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;			
OC	EUTHERIA; PRIMATES.			
RN	[1]			
RC	SEQUENCE FROM N.A.			
RC	TISSUE=PANCREAS;			
RA	MELKONYAN H., CHANG W.C., SHAPIRO J.P., MAHADEVAPPA M., FITZPATRICK P.A.;			
RA	KIEFER M.C., TONEI D.L., UMANSKY S.R.;			
RL	SUBMITTED (SEP-1997) TO EMBL/GENBANK/DBJ DATA BANKS.			
DR	EMBL; AF017988; G2415419; -			
SEQ	SEQUENCE 317 AA; 35577 MW; 2CD69373 CRC32;			
	Query Match 43.5%; Score 901; DB 2; Length 317;			
	Best Local Similarity 45.2%; Pred. No. 9.26e-190;			
	Matches 114; Conservative 58; Mismatches 74; Indels 6; Gaps 4;			
Db	46 SYSKPPQCDIPADPLPCHTVGVKRMVLPNLLHESLAEVVKQOASSWPLLNKRCHSDTQ 105			
QY	14 SY-KRSNCKPIPANLQCHGIEYQNRLPNLLGHETMKVELEQAGAWIPLVMKQCHPDK 72			
Db	106 VFICSLFAPVCLD---RPIYPCRSLSCEAVRAGCAPLMEAYGFFWPEMLHCKHKEFLDNCLC 162			

Qy 142 LPA 144

Search completed: Tue Aug. 4 10:44:04 1998

Job time : 85 secs.

THIS PAGE BLANK (user)

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (C) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

(TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (C) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPsrch_nn n.a. - n.a. database search, using Smith-Waterman algorithm

Run on: Wed Aug 5 03:42:10 1998; MasPar time 1087.57 Seconds
Tabular output not generated. 1494.907 Million cell updates/sec

Title: >rev-trans-seq-2

Description: (1-885) from translate.seq

Perfect Score: 734

N.A. Sequence: 1 ATGTYNCARGGNCNGGWS.....SNATMNGNAARYTNCARTGY 885

Comp: TACRANGTYCNGGNCNGWS.....SNTANKNTTYRANGTYACR

Scoring table: TABLE default

Gap 6

Nmatch STD : Dbase 0; Query 0

Searched: 532261 seqs, 918536377 bases x 2

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database:

emb155

1:em_ba 2:em_htg 3:em_huml 4:em_hum2 5:em_in 6:em_om

7:em_or 8:em_ov 9:em_pat 10:em_pl 11:em_ro

genbank107

12:gb_ba 13:gb_htg 14:gb_in 15:gb_om 16:gb_ov 17:gb_pat

18:gb_ph 19:gb_pl 20:gb_pri 21:gb_pr2 22:gb_ro 23:gb_st

24:gb_sts 25:gb_sy 26:gb_un 27:gb_v1

Statistics: Mean 14.843; Variance 13.039; scale 1.138

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	722	98.4	2028	22	MMU8567 Mus musculus secreted	0.00e+00
2	722	98.4	2031	22	AF017989 Mus musculus secreted	0.00e+00
3	720	98.1	1776	22	D50462 Mouse SDF5 mRNA, Compl	0.00e+00
4	477	65.0	882	21	AF017986 Homo sapiens secreted	3.23e-200
5	144	19.6	1984	21	AF017988 Homo sapiens secreted	7.81e-44
6	134	18.3	2094	21	AF017987 Homo sapiens secreted	2.20e-39
7	131	17.8	2075	21	AF001900 Homo sapiens secreted	4.69e-38
8	131	17.8	4469	21	AF056087 Homo sapiens secreted	4.69e-38
9	123	16.8	2659	22	MMU8566 Mus musculus secreted	1.56e-34
10	122	16.6	4240	15	BTU85945 Bos taurus frizzled-re	4.29e-34
11	122	16.6	7218	17	I66494 Sequence 14 from paten	4.29e-34
12	99	13.5	2075	16	AF006508 Gallus gallus crescent	3.74e-24
13	89	12.1	2421	21	MMU43321 Mus musculus putative	6.16e-20
14	75	10.2	2334	20	HSU43318 Human putative transme	3.50e-14
15	69	9.4	1881	22	AF054623 Mus musculus frizzled-	8.77e-12

16	64	8.7	1770	14	CEU43316	Caenorhabditis elegans	8.02e-10
17	63	8.6	1181	16	AF059570	Xenopus laevis secretate	1.96e-09
18	63	8.6	2831	22	MMU43317	Mus musculus putative	1.96e-09
19	62	8.4	2184	21	HSU82169	Human frizzled homolog	4.76e-09
20	61	8.3	1727	16	AF031831	Gallus gallus 7-transm	1.15e-08
21	61	8.3	2085	14	DMU65589	Drosophila melanogaste	1.15e-08
22	61	8.3	2328	16	AF031830	Gallus gallus 7-transm	1.15e-08
23	59	8.0	2260	22	MMU43320	Mus musculus putative	6.67e-08
24	56	7.6	1913	14	CEU63557	Caenorhabditis elegans	8.99e-07
25	56	7.6	4540	22	RATFRZ2H	Rattus norvegicus Dros	8.99e-07
26	54	7.4	1923	21	HUMFRIZ	Human frizzled gene pr	4.97e-06
27	53	7.2	215	17	I28278	Sequence 5 from paten	1.16e-05
28	52	7.1	1845	14	AF013953	Caenorhabditis elegans	2.69e-05
29	52	7.1	1912	22	RATFRZ2H	Rattus norvegicus Dros	2.69e-05
30	52	7.1	10772	14	AF012089	Drosophila melanogaste	2.69e-05
31	52	7.1	34576	14	CET23D8	Caenorhabditis elegans	2.69e-05
32	50	6.8	10772	14	AF012089	Drosophila melanogaste	1.43e-04
33	49	6.7	565	17	E04076	gDNA encoding envelope	3.25e-04
34	47	6.4	215	17	I28278	Sequence 5 from paten	1.66e-03
35	46	6.3	2624	22	MMU43205	Mus musculus frizzled-	3.72e-03
36	46	6.3	7218	17	I66494	Sequence 14 from paten	3.72e-03
37	45	6.1	25700	14	CELF27E11	Caenorhabditis elegans	8.25e-03
38	44	6.0	354	14	OFU89259	Oxytricha fallax 57kd	1.82e-02
39	43	5.9	354	14	OFU89259	Oxytricha fallax 57kd	3.98e-02
40	43	5.9	565	17	E04076	gDNA encoding envelope	3.98e-02
41	43	5.9	16741	14	PFSC03095	Plasmodium falciparum	3.98e-02
42	43	5.9	95908	14	PFSC03010	Plasmodium falciparum	3.98e-02
43	40	5.4	370	14	OFU89262	Oxytricha fallax 57kd	3.97e-01
44	40	5.4	2478	9	A63310	Sequence 3 from paten	3.97e-01
45	40	5.4	3342	21	AB012911	Homo sapiens mRNA for	3.97e-01

ALIGNMENTS

RESULT	1	MMU8567	2028 bp	mRNA	ROD	22-APR-1997
LOCUS						
DEFINITION		Mus musculus secreted frizzled related protein sFRP-2 (Sfrp2) mRNA,				
ACCESSION		U88567				
NID		91946342				
KEYWORDS		house mouse.				
SOURCE		Mus musculus				
ORGANISM		Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;				
REFERENCE		1 (bases 1 to 2028)				
AUTHORS		Rattner, A., Hsieh, J.-C., Smallwood, P.M., Gilbert, D.J.,				
		Copeland, N.G., Jenkins, N.A. and Nathans, J.				
TITLE		A family of secreted proteins contains homology to the				
JOURNAL		cysteine-rich ligand-binding domain of frizzled receptors				
MEDLINE		Proc. Natl. Acad. Sci. U.S.A. 94 (7), 2859-2863 (1997)				
REFERENCE		2 (bases 1 to 2028)				
AUTHORS		Rattner, A., Hsieh, J.-C., Smallwood, P.M., Gilbert, D.J.,				
		Copeland, N.G., Jenkins, N.A. and Nathans, J.				
TITLE		Submitted (05-FEB-1997) Molecular Biology and Genetics, HHMI/Johns				
JOURNAL		Hopkins University School of Medicine, 725 North Wolfe Street, PC1B				
		805, Baltimore, MD 21205				
FEATURES		Location/Qualifiers				
Source		1. .2028				
		/organism="Mus musculus"				
		/db_xref="taxon:10090"				
		/chromosome="3"				
		/map="near Fgg"				
		236. .1123				
gene		/gene="sfrp2"				
		236. .1123				
CDS		/gene="sfrp2"				
		/codon_start=1				
		/product="secreted frizzled related protein sFRP-2"				
		/db_xref="PID:gl946343"				

[illegible]

RESULT	2	AF017989	2031 bp	mRNA	ROD	21-SEP-1997
LOCUS		Mus musculus secreted apoptosis related protein 1 (Sarpl) mRNA,				
DEFINITION		complete cds.				
ACCESSION		AF017989				
NID		92415420				
KEYWORDS		house mouse.				
SOURCE		Mus musculus				
ORGANISM		Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;				
REFERENCE		Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
AUTHORS		1 (bases 1 to 2031)				
TITLE		Fitzpatrick,P.A., Chang,W.C., Shapiro,J.P., Mahadevappa,M.,				
JOURNAL		SARPs - a new family of proteins that regulate apoptosis				
REFERENCE		Unpublished				
AUTHORS		2 (bases 1 to 2031)				
TITLE		Melkonyan,H., Prochazka,V. and Umansky,S.R.				
JOURNAL		Direct Submission				
FEATURES		Submitted (11-AUG-1997) LXR Biotechnology, Inc., 1401 Marina Way				
source		South, Richmond, CA 94804, USA				
		Location/Qualifiers				
		1..2031				
		/organism="Mus musculus"				
		/db_xref="taxon:10090"				
		/cell_line="10T1/2 quiescent cells"				
		/dev_stage="embryo"				
gene		1..2031				
		/gene="Sarpl"				
CDS		254..1141				
		/gene="Sarpl"				
		/note="similar to frizzled-like receptor CRD"				
		/codon_start=1				
		/product="secreted apoptosis related protein 1"				
		/db_xref="PID:M92415421"				
		/translation="MPRGASLALLVLASHCCLSGARGLFQGPDPFSYKRTNCKPVP				
		ANQLCHGETIQONMLVGLHETMKVEPLFQAGAWIPLVMKQHPOTKKFLCSLAPV				
		CLDDDETIOPHSLCVQVKDRCAPVMSAFGFPWDMCEDRFDPQNDLCIPLASDSDH				
		LLPATPEAKVCACECTKNEDNDIMETLKDKNFALKIKVKEITYLNRPDKIILETKS				
		KTIYKLNGSERDLKLSYLWLKSLQCTCEMDINDINAPILVMQGKGGLVIIVSKRM				
		KQKGRFKRISRSIRKLQC"				
BASE COUNT		460 a	591 c	485 g	495 t	
ORIGIN						
		Query Match	98.4%;	Score 722;	DB 22;	Length 2031;
		Best Local Similarity	58.2%;	Pred. No. 0.00e+00;		
		Matches	515;	Conservative	213;	Mismatches 157; Indels 0; Gaps 0;
Db	254	ATGCCGGGGGCCTGCCTCGCTGCTGTACTGTCTCCTGCCTCGCATGCTGCTGGCG	313			
Qy	1	ATGYTNCAAGNCNGGNWSNYNTYNTYNTTYTTCGNCWSNCAYTGTYTTNGGN	60			
Db	314	TCGGCGGTGGCTCTCCCTTCGGCCAGCCGACACTCTCCTACAGCCGACCACATGC	373			
Qy	61	WSNGCMNGNNYNTTYTTCGNCRCNGAYTTYTSNTAYAAARMGNSNAATGY	120			
Db	374	AAGCCCATTCCCGCCAACCTGCAGCTGTGCCAGCGCATCGAGTACCAAGAACATCGCGCTG	433			
Qy	121	AARCNATCCNGCNAAYTYNCARYTGTGTCAYGSNATNGARTAYCARAAYATGMGYTN	180			
Db	434	CCCAACCTGCTGGCCACGAGACCATCAAGAGAGTGCTGGAGACGGCGGCGCTGGAAT	493			
Qy	181	CCNAAYTYNTNGNCAYGARACNATGAARCATNTNGARCARGNGGNGCNTGGATN	240			
Db	494	CCGCTGTGTCATGAAGTAGTGCACCCGACACAAGAAGTTCCGTGCTCTCTCGCC	553			
Qy	241	CCNTNGTATGAARCATGYCACCGNAYACNARAARTYYTNTGWSNYNTTTCGN	300			
Db	554	CCTGTCTGCTCGACGACCTAGATGAGACCATCCAGCGCGTGTCACTCGCTCTCGGTGCA	613			
Qy	301	CCNGTNTGYTNGAYYTNGAYGARACNATCARCNTGYCAYWSNTNTGYGTNCAR	360			

```
Db 614 GTGAAGGACCGCTGGCCCGCGTATGTCGGCTTCGGCTTCCCTGGCCAGACATGCTG 673
Qy 361 GTNARGAYMGNTGVCNCCNTGATGWSNGCNTYGGNTTCCNTGGCCGAYATGYN 420
Db 674 GAGTCGACCGCTTCCCGAGGACACAGACCTCTGCTATCCCTCGCTAGTAGGACAC 733
Qy 421 GARTGYGAYMGNTTCCNCARGAAYGAYTNTGYATNCCNTGNCNWSNNGAYCAY 480
Db 734 CTCCTGCGCGCCACAGAGAAGCTCCCAAGGTGTGTGAAGCTTGAACACCAAGATGAG 793
Qy 481 YTNYNCCNCGNACNARGACNCCNNAARGTNTGYGARGCNTGYAARAAYARAAYGAY 540
Db 794 GACGACACGACATCATGATAAACCCTTTCTGAAAATGACTTCGCCTGAAATCAAGTG 853
Qy 541 GAYGAYAYGATNATGARGACNTNTGAYARAAYGAYTTCYCNYNARATNARNTN 600
Db 854 AAGGAGATTAACGTATACATCAACAGACACCAAGATCATCTCGAGACAAGAGACCC 913
Qy 601 AARGARATNACNTATATNAAYMGNGAYACNAARATNATNTNGARACNAARNAACN 660
Db 914 ATTACAACTGAACGGCTGTCGGAAGGACCTGAAGAAATCGTGTGGCTCAAA 973
Qy 661 ATNTAAYATTAAYGNGTNSNGARMNGAYITNAARAARWSNGTNTTGGYTNAR 720
Db 974 GACGCTCGAGTGCACCTGTGAGGAGATGAACGACATCAAGCCTCGCTATCTGTCATG 1033
Qy 721 GAYWSYNTNARTGYACNTGYGARGARATGAAYGAYATNAAYGCNCCNTAYTNGTNG 780
Db 1034 GGACAGAACGAGCGGCGGAACTGTGATCACTCGCTGAAACGGTGGCAGAGGCCAG 1093
Qy 781 GGNCAARACRGNGNGARNTNGTNGTNGTNGTNGTNGTNGTNGTNGTNGTNGTNGTNG 840
Db 1094 AGAGAGTTCAAGCGCATCTCCGCGACGATCCGCAAGCTCATGC 1138
Qy 841 MNGARTTYAARMGNATNSNMGNSNATNMGNATNMGNAARTNARTGY 885

RESULT
LOCUS D50462 1776 bp mRNA ROD 24-DEC-1996
DEFINITION Mouse SDF5 mRNA, complete cds.
ACCESSION D50462
NID g1747301
KEYWORDS SDF5.
SOURCE Mus musculus cell_line:ST-2 cdna to mRNA, clone_lib:phage (lgt22a)
ORGANISM Mus musculus
REFERENCE
AUTHORS Shirozu,M., Tada,H., Tashiro,K., Nakamura,T., Lopez,N.D.,
Nazarea,M., Hamada,T., Sato,T., Nakano,T. and Honjo,T.
TITLE Characterization of novel secreted and membrane proteins isolated
by the signal sequence trap method
JOURNAL Genomics 37 (3), 273-280 (1996)
MEDLINE 97092876
REFERENCE
AUTHORS Shirozu,M., Tada,H., Nakamura,T., Nelson,L.D., Martina,N.,
Hamada,T., Sato,T., Tashiro,K., Nakano,T. and Honjo,T.
TITLE Isolation of novel genes encoding for secreted or membrane proteins
using signal sequence trap
JOURNAL Unpublished (1995)
REFERENCE
AUTHORS Shirozu,M.
TITLE Direct Submission
JOURNAL Submitted (01-MAY-1995) to the DDBJ/EMBL/GenBank databases. Michio
Shirozu, Kyoto University, Faculty of Medicine, Department of
Medical Chemistry, Yoshida, Sakyo-ku, Kyoto, Kyoto 606, Japan
(E-mail:kondo@viru1.virus.kyoto-u.ac.jp, tel:81-75-753-4377,
Fax:81-75-753-4388)
FEATURES
Location/Qualifiers
1..1776
/organism="Mus musculus"
```

```

/db_xref="taxon:10090"
/cell_line="ST-2"
/clone_lib="phage (lgt22a) library, ST-2 directional 1"
14..901
/gene="SDF5"
14..901
/gene="SDF5"
/codon_start=1
/db_xref="PID:d1009688"
/db_xref="PID:g1747302"
/translation="MPGPAASLLLLVLAHHCCGASGLFLFGQPDFSYKRSNCKPVP
ANQLCHGIEYQNNRLNLLGHETVKEVLEOAGAWIPLVMKQHPDKKELCSLFAPV
LLDDETIQFCHSLQKQVDRCAPVMSAFGFPWMDLECDREFQDDLDLCPLASSDH
LLDPATEAPKVCCEACKTNEEDNDIMETLCKNDFALKIKVEITIIYINRDKIILETKS
KTIYKLVGVSERDLKSVLMLKLSLOCTCEEMNDINAPYLVMGKQGGELVITSVKRW
OKGOREFKRISRSIRKLCQ"
BASE COUNT 433 a 483 c 405 g 455 t
ORIGIN
Query Match 98.1%; Score 720; DB 22; Length 1776;
Best Local Similarity 58.1%; Pred. No. 0.00e+00;
Matches 514; Conservative 213; Mismatches 158; Indels 0; Gaps 0;
Db 14 ATGCCGCGGGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 73
Qy 1 ATGYTCNARGGNCNGNWSNNTYNTNTNTNTNTNTNTNTNTNTNTNTNTNTNTNTNTNT 60
Db 74 TCGGCGCGTGGCTCTTCTCTTCCGCCAGCCGACCTTCTCTTACAGCGCAGCACTGC 133
Qy 61 WSNCGCMNGNGNTTNTTNTTNGGNCARCCNGAYTTTWSNTAYAAARMGNSNAYTGY 120
Db 134 AAGCCCATCCCGCCCAACTGCGCTGTGCCAGCGCTGCTGCTGCTGCTGCTGCTGCTG 193
Qy 121 AARCCNATNCCNCGNAYTNCARTNTGYCAYGNARTNGARTAYCARAAYATGNGNTN 180
Db 194 CCCAACCTGCTGGGCCACGACCATGAAGGAGTGTGCTGGAGCAGCGCGGCGCTGGATT 253
Qy 181 CCNAAAYTNTNGGNCAYGARACNATGAARGNTNYNGARCGNCGNGCNGTGGATN 240
Db 254 CGCTGTGTCATGAACGAGTGCACCCGCGACACACCAAGAGTTCCTGCTGCTGCTGCC 313
Qy 241 CCNTNGTNGTGAARCACTGTCAYCCNGAYACNAARAARTTYTNTGYWSNTNTTYG 300
Db 314 CCGTGTGCTGCTGCGACGCTAGATGAGACCATCCAGCGCTGTGCTGCTGCTGCTGCTG 373
Qy 301 CCNGTNTGYTNGAYGAYTNGAYGARACNATNCARCCNTGYCAYWSNTNTGYTNCAR 360
Db 374 GTGAAGGACCGCTGCGCCCGGCTGATGTCGCTTCCCTGCTGCTGCTGCTGCTGCTG 433
Qy 361 GTNARGAYMGNTGYCNCNGTNGTNGTNGTNGTNGTNGTNGTNGTNGTNGTNGTNG 420
Db 434 GAGTCGACCGCTTCCCGAGGACACACGACCTCTGCATCCCTCGCTAGTAGGACACAC 493
Qy 421 GARTGYGAYMGNTTCCNCARGAAYGAYTNTGYATNCCNTGNCNWSNNGAYCAY 480
Db 494 CTCCTGCGCGCCACAGAGAAGCTCCCAAGGTGTGTGAAGCTTGAACACCAAGATGAG 553
Qy 481 YTNYNCCNCGNACNARGACNCCNNAARGTNTGYGARGCNTGYAARAAYARAAYGAY 540
Db 554 GACGACACGACATCATGAAACCCCTTTGTAAAATGACTTCGCTGCTGCTGCTGCTGCT 613
Qy 541 GAYGAYAYGAYATNATGARGACNTNTGYAARAAYGAYTTCYCNYNARATNARNTN 600
Db 614 AAGGAGATTAACGTATACATCAACAGACACCAAGATCATCTTGGAGACAAGAGACCC 673
Qy 601 AARGARATNACNTATATNAAYMGNGAYACNAARATNATNTNGARACNAARNAACN 660
Db 674 ATTTCAAGCTGAACGGCGTGTCCGAAAGGACCTGAAGAAATCCGCTGCTGCTGCTCAA 733
Qy 661 ATNTAAYATTAAYGNGTNSNGARMNGAYITNAARAARWSNGTNTTGGYTNAR 720
Db 734 GACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 793
Qy 721 GAYWSYNTNARTGYACNTGYGARGARATGAAYGAYATNAAYGCNCCNTAYTNGTNG
```



```

QY 696 NAARAARWSNGTNTGGYTAARAYWSNYTNCARTGYACNTGYGARGARATGAAYGA 755
Db 1091 CCTCAGCCACCACITCTCATCATGGCCCGCAAGGTGAAGAGCCAGTCTGTCGACGGC 1150
QY 756 YATNAAYGNCNTAYTNGTATGGGNCARAARCARCGGNGGNGARYTNGTATNACNWS 815
Db 1151 CATCCCAAGTGAGCAAAAAACAAGGAGTTCACAAACCTTCATGAAGAAAAATGAAAA 1210
QY 816 NGTNAARMGNTGGCARAARGNCARMGNARTTYAARMGNATNSNMGNWSNATNMGNAA 875
Db 1211 CCATGAGTGC 1220
QY 876 RYTNCARTGY 885

RESULT 7
LOCUS AF001900 2075 bp mRNA PRI 25-JUN-1997
DEFINITION Homo sapiens secreted frizzled-related protein mRNA, complete cds.
ACCESSION AF001900
NID G2213818
KEYWORDS human.
ORGANISM Homo sapiens
Eukaryotes; mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homiidae;
Homo.
REFERENCE 1 (bases 1 to 2075)
AUTHORS Finch,P.W., He,X., Kelley,M.J., Uren,A., Schaudies,R.P.,
Popescu,N.C., Rudikoff,S., Aaronson,S.A., Varnus,H.B. and
Rubin,J.S.
TITLE Purification and Molecular Cloning of a Secreted, Frizzled-Related
Antagonist of Wnt Action
JOURNAL Proc. Natl. Acad. Sci. U.S.A. (1997) In press
REFERENCE 2 (bases 1 to 2075)
AUTHORS Finch,P.W., Aaronson,S.A. and Rubin,J.S.
JOURNAL Direct Submission
Submitted (30-APR-1997) LCMB, DBS/NCI, 9000 Rockville Pike,
Bethesda, MD 20892, USA
REFERENCE 3 (bases 1 to 2075)
AUTHORS Finch,P.W., Aaronson,S.A. and Rubin,J.S.
JOURNAL Direct Submission
Submitted (09-JUN-1997) LCMB, DBS/NCI, 9000 Rockville Pike,
Bethesda, MD 20892, USA
REMARK Sequence update by submitter
FEATURES
    Location/Qualifiers
        1..2075
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /chromosome="8"
            /map="8p11.1-p12"
            /clone="HS1"
            /dev_stage="embryonic"
            /tissue_type="lung"
            /cell_line="M426"
            /cell_type="fibroblast"
            303..1244
                /note="FRP"
            /codon_start=1
            /product="secreted frizzled-related protein"
            /db_xref="PID:G2213819"
            /translation="MGIGRSEGGRRGALGVILGAALLAVGASAEYDVVTSQSDIGP
YQSGRYTPPCQVDIPADRLCHNVYKVMVLPNLEHETMAEYKQASVWPLLNK
NCHAGTFLCLSDIPVCLDRPIYPCRWLCEAVRDSCEPMQFFGFYWPMLKCDKFP
EGDVCITAPPNATEASKPGQTVPCPCONELKSEAIIEHLGASFEALRMKIEVKKE
NGDKITVPKKPLKPIKPKDKDLKLVLYLKNAGDCPCQDLNLSHFLINGRVKS
QYLLTAIHAKWDKNKEFKMKMKMKHNECPFTQSVFK"
BASE COUNT 473 a 596 c 626 g 380 t
ORIGIN

Query Match 17.8%; Score 131; DB 21; Length 2075;
Best Local Similarity 39.4%; Pred. No. 4.69e-38;
Matches 311; Conservative 139; Mismatches 319; Indels 21; Gaps 14;

```

```

Db 449 CTTCTACCAACCCACCTCAGTGCCTGACATCCCGCGGACCTCGGCTGTGCCACAA 508
QY 96 YTTWNTAYAARWGWSNAAYTGYAARCCNATNCNCNAAYYTNCARYTNTGYCAYGG 155
Db 509 CGTGGCTACAAGAAGATGGTGTGCTGCCAACCTGCTGGAGCAGGAGACCATCGCGAGGT 568
QY 156 NATNGARTAYCARAAATATGMNTNCCNAAYYTNYTNGGNCAYGARACNATGAARGART 215
Db 569 GAAGCAGCAGGCGAGCAGTGGTGGTCCCTGCTCAACAAGAATGCCACCGCGGACCCA 628
QY 216 NYTNGARCARCGNGGNCNTGGATNCCNTNGTNGTNGTNGTNGTNGTNGTNGTNGTNG 275
Db 629 GGTCTTCCTCTGCTGCTCTTCGCGCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 679
QY 276 RAARTTYTNTGYWSNYTNTTYGCNCCNGTNTGYTNGAYGAYTNGAYGARACNATNCA 335
Db 680 CCGGTGTGCTGCTGCTGCGAGCGCTGCGGACTCGTGGAGCCGGTCTGTCAGTTCCT 739
QY 336 RCCNTGYCAWWSNYTNTGYTNCARGTNAARGAYMNTGYCNCNGTNGTNGTNGTNGTNG 395
Db 740 CGGCTTCTACTGCGCGAGATGCTTAAGTGTGACAGTTCCTCCGAGGGGACG-TCTGCA 798
QY 396 YGNTTYCMTGCGGNGAYATGTYNGARTGYGAYMNTTYCCNARGAYAYGAYTNTG 455
Db 799 TCGCCATGAGCGCGCCCAATGCCACCGAA-GCCTCCAAGCCCAAGGCAC-AAACGGTGTG 856
QY 456 YATNCCNYTNGCWSNWSNGAYCAYTNTNCCNGCNCARGARGCNCNARGTNTG 515
Db 857 TCCTCCTGTGACAACGAGT-TGAA-ATC-TGAGGCCATCATGAACATCTCTGTGCCAG 913
QY 516 YGARGCNTGYAARAAYAAARAYGAYGAYAYATNATGGARACNYTNTGYAARA 575
Db 914 CGAGTTTGCCTGAGCATGAARATAAAGAGTGAARAAAGAAATGGCGACAGAAGAT 973
QY 576 YGATTYTGNNTNAARTNAARGATNACNTATATNAAYMNGVNGAYACNARAT 635
Db 974 TGTCCCAAGA--AGAAG-AAGCCCTG---AAGTTGGGCCCCCATCAAGAAGAAGGACCT 1027
QY 636 NATYNTGARACNAARWSNAARACNATNTAYAARYTNAAYGNGTNGSNGARMNGAYYT 695
Db 1028 GAAGAAGCTTGTCTCTACCTGAAGAATGGGCTGACTCTCCCTGCCACAGCTGGACAA 1087
QY 696 NAARAARWSNGTNTGYTNGTNAARGAYWSNTNCARTGYACNTGYGARGARATGAAYGA 755
Db 1088 CPTCAGCCACCACTTCTCATCATGCGCGCAGGTTGAAGAGCCAGTCTGCTGACGCG 1147
QY 756 YATNAAYGNCNTAYTNGTNGTNGTNGTNGTNGTNGTNGTNGTNGTNGTNGTNGTNG 815
Db 1148 CATCCCAAGTGGCAAGAAAAACAAGGAGTTCAAAACTTCATGAAGAAAAATGAAAA 1207
QY 816 NGTNAARMGNTGGCARAARGGNCARMGNCARTTYAARMGNATNSNMGNWSNATNMGNAA 875
Db 1208 CCATGAGTGC 1217
QY 876 RYTNCARTGY 885

RESULT 8
LOCUS AF056087 4469 bp mRNA PRI 08-APR-1998
DEFINITION Homo sapiens secreted frizzled related protein mRNA, complete cds.
ACCESSION AF056087
NID G3033550
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 4469)
AUTHORS Zhou,Z. and Wang,J.
TITLE Upregulation of human secreted Frizzled homologue in apoptosis and
its down regulation in breast tumors
JOURNAL Unpublished

```

```
REFERENCE 2 (bases 1 to 4469)
AUTHORS Zhou,Z. and Wang,J.
TITLE Direct Submission
JOURNAL Submitted (29-MAR-1998) Radiohemmet Research Laboratory, CKK,
Karolinska Institute, Solnavagen, Stockholm S-171 76, Sweden
FEATURES
source
1..4469
/organism="Homo sapiens"
/db_xref="taxon:9606"
303..1244
/function="putative Wnt antagonist"
/note="FRP; frizzled homolog; cell cycle and apoptosis
related; down regulated in breast tumorigenesis"
/codon_start=1
/product="secreted frizzled related protein"
/db_xref="PID:g303351"
/translation="MGIRSGRRGALGVLLAALLAVGSASEYDVYVFSQSDIG
YOSGRFYTPQCVDPADLRICHNVGYKKVLPNLEHETMAEVKQQAASWVPLLN
NCHAGTVFLCSLFAVDRPIPCRWLCEAVRDSCEPVMOFFGYFPEMLKCDKFP
EGDVCIAMTPPNATEASKPOGTTCPCDNLKSEALIEHLCASEFALRMKIEVKKE
NGDKKIVPKKKPLKGLPKKKDLKLVLYLKGADCPCHOLDNLSHFLIMGRKVK
QYLLTAIHWDKKKEKFNKMKKNHECPTFSQVFX"
BASE COUNT 1121 a 1097 c 1120 g 1131 t
ORIGIN
Query Match 17.8%; Score 131; DB 21; Length 4469;
Best Local Similarity 39.4%; Pred. No. 4,69e-38;
Matches 311; Conservative 139; Mismatches 319; Indels 21; Gaps 14;
Db 449 CTTCTACACCAAGCCACCTCAGTGGTGGACATCCCGCGGACCTGCGGCTGTGCCACAA 508
QY :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:
96 YTTYSNTAYAAARMGNWSNAAYTGAARCCNATNCCGNCNAAAYTNCARTYNTGYCAYGG 155
Db 509 CGTGGCTACAGAGATGGTCTGCCCAACCTGCTGGAGCAGACGACCATCGCGGAGGT 568
QY :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:
156 NATNGARTAYCARAAYATGMGNTYNTCCNAAAYTNTGNCNCAYGARACNATGAARGGT 215
Db 569 GAAGCAGCAGGCGAGCAGCTGGTGGTCCCTGCTCAACAAGAACTGCACCGCGGGACCCA 628
QY :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:
216 NYTNGARCARGCNGGNCNTGTGATNCCNTNGTNGTNGATGAARCATGYCAYCCNGAYACNAA 275
Db 629 GGTCTTCCTCTGCTCGCTTTCGCGCCGCTGTCGCT-G--GACC-GG-C-----CCATCTA 679
QY :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:
276 RAARTTYTNTGYWSNYTNTTYGCNCCGNTGTGYTNGAYGAYTNGAYGARACNATNCA 335
Db 680 CCGGTGCTGCTGCTGCGAGCGGCTGCGGACCTGCTGCGAGCGGTCATCGAGTCTT 739
QY :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:
336 RCNTGYCAYWSNYTNTGYTNCARGTNAARGAYMGNTGYGCNCCNGTNGTNGSNCNTT 395
Db 740 CGGCTTCTACTGCGCGGAGATGCTTAAGTGTGACAAAGTTCCTCCGGAGGGGACG-TCTGCA 798
QY :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:
396 YGNTTYCCTGCGCNGAYATGTNGARTGYGAYMGNTTYCCNCARGAYAAAYGAYTNTG 455
Db 799 TCGCCATGACGCGCCCAATGCCACCGAA-GCCTCCAAGCCCCCAAGGCAC-AACGGTGTG 856
QY :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:
456 YATNCCNTNGCNSWSNGAYCAVYNTNTCCGNCNACNGARGARGCCNCCNAARGTNTG 515
Db 857 TCCTCCTGTGACAGGAGT-TGAA-ATC-TGAGGCCATCATGACATCTCTGTGCCAG 913
QY :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:
516 YGARGCNTGYAARAAYAAAYGAYGAYAAAYATNATGARGACNYTNTGYAARA 575
Db 914 CGAGTTTGACATGAGGATAAATAAGAAAGTGAAGAAAGAAATGGCGCAAGAAAGT 973
QY :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:
576 YGAYTYGCTNTAARATNAAGTNAARGATNACNTATYATNAYMGNGATACNAARAT 635
Db 974 TGTCCCAAGA--AGAA--AAGCCCTG---AAGTTGGGCGCCATCAAGAAAGGACCT 1027
QY :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:
636 NATNTNGARACNAARWSNAARACNATNTAYAARYTNAAYGNGINGNWSNGARMNGAYT 695
Db 1028 GAAGAGCTTGTGCTGTACCTGAAGAAATGGGCGCTGACTGCTCCCTGCCACCACTGGACAA 1087
QY :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:
696 NAARAARWSNGTNTYNTGGYTNAAARGAYWSNYTNCARTGYACNTGYGARGARATGAYGA 755
```

```
Db 1088 CUTCAGCCACCACCTTCCTCATCATGGCCGCAAGGTGAAGACCCAGTACTTCTGCACGCG 1147
QY :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:
756 YATNAAYCNCNTAYTINGTATGGTGGNCARAAACARGCNGGNGARYTINGTATNACNWS 815
Db 1148 CATCCACAAGTGGGACAGAAAACACAGAGTTCACAAACTTCATGAAGAAAATGAAAAA 1207
QY :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:
816 NCTNAARMGNTGGCARAARGGNCARMGNGARTTYAARMGNATNWSNMGNSNATNMGNA 875
Db 1208 CCATGAGTGC 1217
QY :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:
876 RYTNCARTGY 885
RESULT 9
LOCUS MMU88566 2659 bp mRNA ROD 22-APR-1997
DEFINITION Mus musculus secreted frizzled related protein sFRP-1 (Sfrpl) mRNA,
complete cds.
ACCESSION U88566
NID 91946340
KEYWORDS house mouse.
SOURCE Mus musculus
ORGANISM
Eukaryota; Eukaryota; Mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Mammalia; Euthera; Rodentia; Sciurognathi; Muridae;
Murinae; Mus.
REFERENCE 1 (bases 1 to 2659)
AUTHORS Rattner,A., Hsieh,J.-C., Smallwood,P.M., Gilbert,D.J.,
Copeland,N.G., Jenkins,N.A. and Nathans,J.
TITLE A family of secreted proteins contains homology to the
cysteine-rich ligand-binding domain of frizzled receptors
Proc. Natl. Acad. Sci. U.S.A. 94 (7), 2859-2863 (1997)
MEDLINE 97250455
REFERENCE 2 (bases 1 to 2659)
AUTHORS Rattner,A., Hsieh,J.-C., Smallwood,P.M., Gilbert,D.J.,
Copeland,N.G., Jenkins,N.A. and Nathans,J.
TITLE Direct Submission
JOURNAL Submitted (05-FEB-1997) Molecular Biology and Genetics, HHMI/Johns
Hopkins University School of Medicine, 725 North Wolfe Street, PCTB
805, Baltimore, MD 21205
FEATURES
source
1..2659
/organism="Mus musculus"
/db_xref="taxon:10090"
/chromosome="8"
/map="between Plat and Fgf1"
257..1201
/gene="Sfrpl"
257..1201
/gene="Sfrpl"
/codon_start=1
/product="secreted frizzled related protein sFRP-1"
/db_xref="PID:g1946341"
/translation="MGVRSARGGGAAGVLLAALLAALAAASASDYVYVFSQSDIG
SYSGREYTKPPQCVDPADLRICHNVGYKKVLPNLEHETMAEVKQQAASWVPLLN
KNCHAGTVFLCSLFAVDRPIPCRWLCEAVRDSCEPVMOFFGYFPEMLKCDKFX
PEGDVCIAMTPPNATEASKPOGTTCPCDNLKSEALIEHLCASEFALRMKIEVKKE
NGDKKIVPKKKPLKGLPKKKDLKLVLYLKGADCPCHOLDNLSHFLIMGRKVK
QYLLTAIHWDKKKEKFNKMKKNHECPTFSQVFX"
BASE COUNT 645 a 738 c 719 g 553 t 4 others
ORIGIN
Query Match 16.8%; Score 123; DB 22; Length 2659;
Best Local Similarity 38.9%; Pred. No. 1.56e-34;
Matches 307; Conservative 140; Mismatches 322; Indels 21; Gaps 14;
Db 406 CTTCTACCAAGCCCGCCGACGTGGTGGACATCCCGTGGACCTGAGCTGTGCCACAA 465
QY :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:
96 YTTYSNTAYAAARMGNWSNAAYTGAARCCNATNCCNCAAYTNCARTYNTGYCAYGG 155
Db 466 CGTGGCTACAGAGATGGTGTGCTGCCCAACCTGCTGGAGCAGACGACCATCGGAGAGGT 525
QY :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:
156 NATNGARTAYCARAAYATGMGNTYNTCCNAAAYTNTGNCNCAYGARACNATGAARGGT 215
```

D	b	526	GAAGCAGCAGCGCCACGACTGGGTGGCGCTGCTCAACAAGAATGCCACATGGGCACC	A	585
			: : : :	: : :	
Q	y	216	NYTNGARCANGCGNGCNTGGATNCCNTNGTNATGAARCATGYCAYCNGYACNA	A	275
			: : : :	: : :	
D	b	586	GGTCTTCCTCTGTTCCGCTCTTCGCGCCCGTGCTGCT--G--GACC-GG-C----	C	636
			: : : :	: : :	
Q	y	276	RAARTTYTNTGYSNYTNTTGYNCNGNTNGTYYTNGAYGAYITNGAYGARACNATN	C	335
			: : : :	: : :	
D	b	637	CCCGTTCGCTGGCTGCGAGGCGGTGCGGACTCTGCGGACGCGGTCTATGCAGTCT	T	696
			: : : :	: : :	
Q	y	336	RCCNTGYCAYSNTYTNGTNCARGINAARGAYWMTGYGCCNCGTNTATGNSCN	T	395
			: : : :	: : :	
D	b	697	CGGCTTCTACTGGCCGAGATGCTCAAATGTGACAAGTTCGCCGAGGCGAGC-T	-	753
			: : : :	: : :	
Q	y	396	YGGNTTYCCNTGGCCNGAYATGTYTNGARTGYGAWGNTTYCCNCARGAAYA	A	455
			: : : :	: : :	
D	b	754	CATGCCCATGACCCCGCCCATACACCGGAAGCTCTTAAGCCCCCAAGGTACAC	C	813
			: : : :	: : :	
Q	y	456	YATNCCNTYNGCNWSNNGAYCAYTYNTNCCNGCNARGARGCNCNNA	A	515
			: : : :	: : :	
D	b	814	TCCTCCATGCACAACAGAGT-TGAAGTC-AGAG-GCCATCATTTGAACATCT	C	870
			: : : :	: : :	
Q	y	516	YGARGCNTGYAARAAYARAAYGAYGAYAAVATATGARGACNTYNTGYA	A	575
			: : : :	: : :	
D	b	871	CGAGTTTTGCATCGAGGATGAAAATCAAGAAGTGAAGAAGAAAACGGTGACA	A	930
			: : : :	: : :	
Q	y	576	YGAYTTYGCNTAARATNAARGINAARGARATNACNTATATNAAYMNGAYAC	N	635
			: : : :	: : :	
D	b	931	TGTCGCCAAGA--AGAAG--AAACCCCT-TG-AGCTGGGCGCCATCAAGAAG	A	984
			: : : :	: : :	
Q	y	636	NATNYTNGARACNAARWSNARACNATNTAARYTNAAYCGGNGTNWSNGARM	G	695
			: : : :	: : :	
D	b	985	GAAGCGCTTGCTGCTTCTGAGAACGGTGCCGACTGTCCCTGCCACAGCTGC	A	1044
			: : : :	: : :	
Q	y	696	NAARAARWSNNTINTGGYTNAAARGAYSNYNCARTGYACWTGARGARATGA	I	755
			: : : :	: : :	
D	b	1045	CCTCAGCCACAACCTTCTCATCATGGGCCCAAGGTGAAGAGCGCAGTACCT	G	1104
			: : : :	: : :	
Q	y	756	YATNAAYGCNCNTAYTYTNGTNGTATGGCNCARCARCGGNGNGARYTNG	T	815
			: : : :	: : :	
D	b	1105	CATTCAAGTGGGACAGAAAAACACAGGATTCAAAACTTCAAGAGAGATGA	A	1164
			: : : :	: : :	
Q	y	816	NGTNAARMGNTGGCAARAAGGNCARMGNGARTTYAARMGNATNWSNMGN	S	875
			: : : :	: : :	
D	b	1165	CCACGAGTGT 1174		
			: : : :		
Q	y	876	RYTNCARIGY 885		
			: : : :		
<hr/>					
RESULT	10	BTU85945	4240 bp mRNA	MAM	21-AUG-1997
LOCUS		Bos taurus	fizzled-related protein Frza mRNA,		complete cds.
DEFINITION		Accession			
ACCESSION		U85945			
NID		g2337936			
KEYWORDS					
SOURCE		Bos taurus.			
ORGANISM		Bos taurus			
REFERENCE		Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;			
AUTHORS		Vertebrata; Mammalia; Eucheria; Artiodactyla; Ruminantia; Pecora;			
TITLE		Bovidae; Bovidae; Bovinae; Bos.			
JOURNAL		1 (bases 1 to 4240)			
REFERENCE		Duplaa,C. and D'Amore,P.A.			
AUTHORS		Identification and cloning of a novel secreted form of mammalian			
TITLE		fizzled: evidence to suggest a role in the control of growth and			
JOURNAL		differentiation			
FEATURES		Unpublished			
source		2 (bases 1 to 4240)			
		Duplaa,C. and D'Amore,P.A.			
		Direct Submission			
		Submitted (20-JAN-1997) INSERM U441, AV du Haut Leveque, Pessac			
		33600, France			
		Location/Qualifiers			
		1. 4240			

Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata; Vertebrata; Eutheria; Rodentia; Sciurognathi; Myomorpha; Muridae; Murinae; Mus.

REFERENCE
AUTHORS
1 (bases 1 to 2421)
Wang Y., Macke J.P., Abella B.S., Andreasson K., Worley P., Gilbert D.J., Copeland N.G., Jenkins N.A. and Nathans J.
TITLE
A large family of putative transmembrane receptors homologous to the product of the Drosophila tissue polarity gene frizzled
J. Biol. Chem. 271 (8), 4468-4476 (1996)

JOURNAL
MEDLINE
96224032

REFERENCE
AUTHORS
2 (bases 1 to 2421)
Abella B., Wang Y., Macke J.P. and Nathans J.
TITLE
Direct Submission
Submitted (14-DEC-1995) Jeremy Nathans, Molecular Biology and Genetics, Johns Hopkins Medical School, 725 N. Wolfe Street, Baltimore, MD 21205

JOURNAL

FEATURES
source
Location/Qualifiers
1..2421
/organism="Mus musculus"
/db_xref="taxon:10090"
/chromosome="18"
/map="between Tpl2 and Cdh2"
188..2245
/gene="frizzled 8"
188..2245
/note="frizzled 8"
/note="putative transmembrane receptor"
/codon_start=1
/product="transmembrane receptor"
/db_xref="PDB:g1l51260"
/translation="MEMGYLLEVTLSLAALVLRSSGAAAAAKELACOEITVLPLCK
GIGYNTYMPNQNFDAGLEVHQFWLPVEIQSPDLKFLCSMYTICLEDYKK
PLPCRSVCERAKAGAPXOYGFWAMPDRMCDRLPEQGNPDTCMDYNRTDLTAA
PPSPRLPPPPEGPQPSSGGSRPPGARPHRRGGSGSDAAAAPPGRGKGARPP
GGAAACEPGCCQRAPWVSSEHRLYNEVKTGQIANCALPCHNEFPESODRAFVFI
WIGLWSVLCFTFAIVSYFLIDMERFPIPERFIILSACYLFVSVGVLVRLVAGHEK
VACSAGPAGAGRGAGGAAGAAGARGASPGARGEYEELGAEVQHVRVETTPGA
LCVTFVLLVYFFGMASSIMWWIISLTWFLAKMKWGNEAIAGYSQYFHAAWLVPYSK
SIAVLASLVDPGVAGICKVGNOSLDNLRFGLVALPLVIFGTFTWLLAGFYSLFRI
RSVIKQOGPTKTHKEKLIRGLFTLVTPVAANVVACLFEYEOHNRPWEATHSCGP
CLRDLPQARRPDYAFVMLKYPNCVLVVGITTSVWVWSGKTLESRALCTGCCWASKP
AALVAGAGGSGPSGPGGGGGGGGSGLSDYSTGLTWRSSTASSSVSQKMPLL
SQY"

BASE COUNT 403 a 772 c 779 g 467 t

ORIGIN

Query Match 12.1%; Score 89; DB 22; Length 2421;
Best Local Similarity 41.1%; Pred. No. 6.16e-20;
Matches 124; Conservative 47; Mismatches 131; Indels 0; Gaps 0;

Dbb 306 TGCGGTGTGCAAGGATCGGTATACAACACTACACTTACATCCCCAACAGTTCAACCACG 365
| | : : || : | | | | | | | | : | | : | | : | | : | | : | | : | | : |
Qyb 140 TNCARYTNGTCAYGGNATNGARTAYCARAAYATGMNYTNCCNANYTYNTNGNCNAYG 199
| | | | | | | | | | : | | : | | : | | : | | : | | : | | : | | : |
Dbb 366 ACACGCAAGATGAGCGGGCCCTAGAGGTGCACCAAGTTTTGGCCGCTGGTGAGAGATACAGT 425
| | | | | | | | | | : | | : | | : | | : | | : | | : | | : | | : |
Qyb 200 ARACNATGAARGARGTNYTNGARCARGCNGCGNCGTGGATNCCNTNGTNATGAARCAT 259
| | | | | | | | | | : | | : | | : | | : | | : | | : | | : | | : |
Dbb 426 GCTCCCGGACCTCAAGTCTTCTGTGTAGCATGTACACGCCCATCTGCCTGGAGACT 485
| : | | | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : |
Qyb 260 GCYACCNAGYACNAARAARTYYTNTGYWSNYNTTYGCMCCNGTNGTYTNGAYGAY 319
| | | | | | | | | | : | | : | | : | | : | | : | | : | | : | | : |
Dbb 486 ACAAGAGCCCTCGCCGCTGTGCGCTGTGCTGAACGGCGCAAGCGCGCTGCGCGC 545
| | : | | | | | | | | : | | : | | : | | : | | : | | : | | : | | : |
Qyb 320 TNGAYGARACNATWCARCNTGYCAWSNYTNGTGTNCARGTNAARGAYMGNTGYGCNC 379
| | | | | | | | | | : | | : | | : | | : | | : | | : | | : | | : |
Dbb 546 CGCTCATCGGCACATAGCGCTTTCGTTGGCTGACCGCATCGCTGCGATCGGTGCGG 605
| | | | | | | | | | : | | : | | : | | : | | : | | : | | : | | : |
Qyb 380 CNGTNATGWSNGCNTTYGGNTTYCNGTGGCCNGAYATGTYTNGARTGYGAYMGNTTYCCNC 439
| | | | | | | | | | : | | : | | : | | : | | : | | : | | : | | : |

Dbb 606 AG 607
| :
Qyb 440 AR 441

RESULT	14	HSU43318	2334 bp	mrna	PRI	24-FEB-1996
LOCUS		Human putative transmembrane receptor (frizzled 5)	mrna	complete		
DEFINITION		cds.				
ACCESSION		U43318				
NID		g1151251				
KEYWORDS		human.				
SOURCE		Homo sapiens				
ORGANISM		Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata; Vertebrata; Eutheria; Primates; Catarrhini; Homi				
REFERENCE		1 (bases 1 to 2334)				
AUTHORS		Wang,Y., Macke,J.P., Abella,B.S., Andreasson,K., Worley,P., Gilbert,D.J., Copeland,N.G., Jenkins,N.A. and Nathans,J.				
TITLE		A large family of putative transmembrane receptors homologous to the product of the Drosophila tissue polarity gene frizzled				
JOURNAL		J. Biol. Chem. 271 (8), 4468-4476 (1996)				
MEDLINE		96224032				
REFERENCE		2 (bases 1 to 2334)				
AUTHORS		Abella,B., Wang,Y., Macke,J.P. and Nathans,J.				
TITLE		Direct Submission				
JOURNAL		Submitted (14-DEC-1995) Jeremy Nathans, Molecular Biology and Genetics, Johns Hopkins Medical School, 725 N. Wolfe Street, Baltimore, MD 21205				
FEATURES		Location/Qualifiers				
source		1..2334				
		/organism="Homo sapiens"				
		/db_xref="taxon:9606"				
		/chromosome="2"				
		/map="2q33-34"				
		/tissue_type="retina"				
gene		321..2078				
		/gene="frizzled 5"				
CDS		321..2078				
		/gene="frizzled 5"				
		/note="putative transmembrane receptor"				
		/codon_start=1				
		/product="transmembrane receptor"				
		/db_xref="PID:g1151252"				
		/translation="MARPDSPAPSLLLLLLAOLVGPAASAKAPVQELTPMCRGIGYLNTHPNFNDTDEAGLEWHFPLVEIOCSPLREFLCTMTPTCLDPDTHKPEPPCRVCVERAKGCSPLMRQYGFAPWFMSCDLPLVGRDAEVLCDMNRSEATVPAPRPFAKPTLPFGPASGECAGGPFVCKREPFVPIKSHPLNKYKTVGPVNCNPFVQPSFGSADRTATGWLGSVLCFTSTTATFLIDMTFRYPERPIFLSGAAYCVLSVSLGFLVRLVGHASVACSRHNHIVETTGPACTIVLLVYVFGMASSIMCVILSLTWFLAAMKWEATAGYGYFHLAAMLIPSVKSTITALALSSVDGDPVAGICVGNQNLNSRFVGLPIVLVLVLTFLLAGFVSLFRISVTKOGGTTDKLEKLMIRIGFTITLVPAISIVACYLQYHRESWEALFCACPGHDGTPRAKEPIWLMLKLTFCMIVGTSVGTWTSKGVTSWRSRF"SRCCRPFRGHKSGGMAAGDPYEAASALTGRITGPPGPAATYHKVQSLSHV"				
BASE COUNT		356 a	803 c	736 g	439 t	
ORIGIN						
Query Match		10.2%	Score 75;	DB 20;	Length 2334;	
Best Local Similarity		39.7%;	Pred. No. 3.50e-14;			
Matches		116;	Conservative	44;	Mismatches 132;	Indels 0; Gaps 0;
Db	439	TGTCGGCGGCATCGGCTACACCTGAGGCACATGCCCAACCGTTCACACGACACGC	498			
Qy	146	TNTGAYGNGATNGARTAYCARAAYATGMGNYNCCNAAIYTYNTNGNCAYGARACNA	205			
Db	499	AGGACGAGCGGGCGCTGGAGGTGCACCAGTTCTTGGCCGCTGGTGAGATCAATGCTCGC	558			
Qy	206	TGAARGARTNYTNGARCARGCNGGNCNTGGATNCCNTYNTNGTATGAARCAITGYCAYC	265			
Db	559	CGGACCTGCGGCTTCTCTATGCACTATGTACAGCGCCATCTGTCTGCCGACCTACCACA	618			
Qy	266	CNGAYACNAARAARTTYTNTGYSWNTTYTNGNCCNGTNGTYTNGAYGAYNGAIG	325			
Db	619	AGCGGCTGCCCGCTCGCGTGTGCGAGCGCGCAAGCGCGGCTGTCTCGCGCGTGA	678			

```
Qy 326 ARACNATNARCCNTGYCAYNSNYNTGYTNCARGTNAARGAYMNGTGYGCNCCNGTNA 385
Db 679 TCGCGCCAGTACGGCTTCGCCGCGGAGCGCATGCGTGGACGCGCTCC 730
Qy 386 TGWSNGCNTTYGGNTTYCCNTGGCCNGAYATGTYTNGARTGYGAYMNTTYCC 437

RESULT 15
LOCUS AF054623 1881 bp mRNA ROD 08-APR-1998
DEFINITION Mus musculus frizzled-1 mRNA, complete cds.
ACCESSION AF054623
NID 93025755
KEYWORDS house mouse.
SOURCE Mus musculus
ORGANISM Mus musculus
Eukaryota; Chordata; Vertebrata; Mammalia; Eutheria;
Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 1881)
AUTHORS Xu, L., Muragaki, Y., Olsen, B. and Li, Y.
TITLE Direct Submission
JOURNAL Submitted (18-MAR-1998) Cell Biology, Harvard Medical School, 220
Longwood Ave., Boston, MA 02115, USA
FEATURES
source
1..1881
/organism="Mus musculus"
/strain="C57BL"
/db_xref="taxon:10090"
/dev_stage="10.5 day embryo"
1..1881
/note="similar to Drosophila polarity gene"
/gene="frizzled-1"
1..1881
/gene="frizzled-1"
/codon_start=1
/product="Frizzled-1"
/db_xref="PID:g3025756"
/translation="MAEEAPSESRAAGRLSLEALPGRREEVGHEDTASHRRPR
ADPRWASGLLLLEAPLLGLVRAAGQVSGQQAQPPPPQPOQSQYNGERG
ISIPDHGYCOPISPLCTDMAYNOT IMPNLIGHTNOEDAGLEVHQFPLVKVQSABL
KFFLCSSWAPCTVLEQALPCRSICERAROGCEALNKKFGFQHPDILKCEKFPVHGA
GELCVQNTSDKGTPTPSLLPEFWSNGHGGGIRGGYGGAGTVERGFSCPRALR
VPSLYNHFGEKDCGAPCEPTKYVGLMYFGEPLRFSRTWIGVSWSCASTLFTVL
TYLVDPFRFSYPERPIISLSCGYTAVAYIAGFLLEDVYVNCDFEDGARTVAOQT
NKECTTILFMLYFFSMASISWWILSTWFLAAGMKWHEAIEANSQYFHLAAWAVP
AIKTTITIALQGVDPDVLGCVFLGLANNVADLGFVLAPFLVILFICTSELLAGFVSL
FRITIMKHDKTEKLEKLMVRIGVFSVLTVPATVIACFYEQAFRQWERSWVA
QSCSKYAIPCPLHGGGVPHPMPSPDFTVMIKYILMNSWRKFYRLTNSKOGET
TV"
BASE COUNT 337 a 602 c 575 g 367 t
ORIGIN
Query Match 9.4%; Score 69; DB 22; Length 1881;
Best Local Similarity 39.9%; Pred. No. 8.77e-12;
Matches 119; Conservative 48; Mismatches 128; Indels 3; Gaps 3;
Db 347 TCCCGCTGTGCACGACATGGGTACACACGACCATCATGCCAACCTGTGGGCCACA 406
Qy 140 TNCARYTNTGYCAYGNATNGARTAYCARAYATGMGNVNTCCNAAYYTNTNGNCAYG 199
Db 407 CGAATCAGGAGGACCGCGCTGTGGAGGTGCACCATCTTACCCCTCTGGTGAAGTGCAGT 466
Qy 200 ARACNATGAARGARTNYTNGARCARGCNGGNCNTGGATNCCYNTNGTNGATGAARCART 259
Db 467 GCTCCGCGGAGCTCAAGTTCTTCCTGTGCTCATATACGCGCTGTGTCACCG-T-ACT 524
Qy 260 GYCACCGAYACNAARAARTTYTNTGWSNYTNTTYGCNCCNGTNGYTNAYGAY 319
Db 525 -GGACGAGCGGTACCGCTCCGCTCCCTGTGCGAGCGCGCACCGAGGCGTGGAGG 583
Qy 320 TNGAYGARACNATNARCNCNTGYCAYSNYTNTGYTNCARGTNAARGAYMNGTGTCC 379
Db 584 CGCTCATGAACAAGTTCGGCTTCCAGTGCCGACACACTCAAGTGCAGAGAGTTCCC 641
```

Qy 380 CNGTNAATGWSNGCNTTYGGNTTYCCNTGGCCNGAYATGTYTNGARTGYGAYMNTTYCC 437

Search completed: Wed Aug 5 04:00:29 1998
Job time : 1099 secs.

THIS PAGE BLANK (USPTO)

W P R F A

(TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPSrch_nn n.a. - n.a. database search, using Smith-Waterman algorithm
Run on: Wed Aug 5 04:16:55 1998; MasPar time 140.38 Seconds
Tabular output not generated.

Title: >rev-trans-seq-2
Description: (1-885) from translate.seq
Perfect Score: 734
N.A. Sequence: 1 ATGTCNARGCNCNGCNGWS.....SNATMGNAAATYTCARTGY 885
Comp: TACRANGTCNCGNGCNGCNGWS.....SNTAKNCNTTYRANGTYACR

Scoring table:
Gap 6

Nmatch STD : Dbase 0; Query 0

Searched: 188442 seqs, 68026449 bases x 2

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: n-geneseq32
1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7
8:part8 9:part9 10:part10 11:part11 12:part12 13:part13
14:part14 15:part15 16:part16 17:part17 18:part18
19:part19 20:part20 21:part21 22:part22 23:part23
24:part24 25:part25 26:part26 27:part27 28:part28
29:part29 30:part30 31:part31 32:part32 33:part33
34:part34 35:part35 36:part36 37:part37 38:part38
39:part39 40:part40

Statistics: Mean 11.324; Variance 13.136; scale 0.862

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	89	12.1	2421	37	T89892	Human frizzled gene 8
2	75	10.2	2334	37	T89889	Human frizzled gene 5
3	64	8.7	1770	37	T89887	Caenorhabditis frizzl
4	63	8.6	2828	37	T89888	Mouse frizzled gene 4
5	61	8.3	2344	37	T89885	Drosophila frizzled g
6	59	8.0	501	3	N50030	Sequence encoding new
7	59	8.0	2259	37	T89891	Mouse frizzled gene 7
8	56	7.6	501	3	N50031	Sequence encoding new
9	54	7.4	501	3	N50033	Sequence encoding new
10	53	7.2	498	3	N50034	Sequence encoding new
11	53	7.2	501	3	N50028	Sequence encoding new
12	53	7.2	565	6	Q35072	HCV envelope region n
13	52	7.1	501	3	N50029	Sequence encoding new

14	51	6.9	501	3	N50023	Sequence encoding new
15	51	6.9	501	3	N50032	Sequence encoding new
16	50	6.8	501	3	N50023	Sequence encoding new
17	49	6.7	501	3	N50024	Sequence encoding new
18	49	6.7	501	3	N50025	Sequence encoding new
19	49	6.7	984	17	Q94336	Degenerate Alteromona
20	48	6.5	501	3	N50026	Sequence encoding new
21	48	6.5	501	3	N50027	Sequence encoding new
22	47	6.4	498	3	N50034	Sequence encoding new
23	47	6.4	501	3	N50029	Sequence encoding new
24	47	6.4	501	3	N50031	Sequence encoding new
25	47	6.4	501	3	N50032	Sequence encoding new
26	47	6.4	501	3	N50028	Sequence encoding new
27	46	6.3	1302	38	T90224	Magnetospirillum sp.
28	46	6.3	2624	37	T89886	Mouse frizzled gene 3
29	45	6.1	501	3	N50030	Sequence encoding new
30	45	6.1	501	3	N50033	Sequence encoding new
31	45	6.1	501	3	N50027	Sequence encoding new
32	44	6.0	501	3	N50024	Sequence encoding new
33	43	5.9	501	3	N50026	Sequence encoding new
34	42	5.7	172	32	T76363	Human interleukin 8 a
35	42	5.7	204	1	N81164	Base substituted E.co
36	41	5.6	204	1	N81164	Base substituted E.co
37	41	5.6	565	6	Q35072	HCV envelope region n
38	41	5.6	657	7	Q43519	Degenerate FMN reduct
39	41	5.6	1065	17	Q94335	Degenerate Alteromona
40	40	5.4	91	9	O51746	Oligonucleotide probe
41	40	5.4	2478	31	T67287	Soluble starch synthase
42	40	5.4	3871	2	N71302	HSV-1 gB and surround
43	39	5.3	91	9	O51746	Oligonucleotide probe
44	39	5.3	501	3	N50025	Sequence encoding new
45	39	5.3	984	17	Q94336	Degenerate Alteromona

ALIGNMENTS

RESULT 1
ID T89892 standard; DNA; 2421 BP.
AC T89892:
DT 27-APR-1998 (first entry)
DE Mouse frizzled gene 8 (Mfz8) encoding a Wnt receptor.
KW Wnt receptor; mouse frizzled gene 8; Mfz8 gene;
KW signal transduction; cancer; cell growth; cell proliferation;
KW mammary tumour; oncogene; therapy; ds.
OS Mus musculus.
FH Key Location/Qualifiers
FT CDS 188..2245
FT /tag= a
PN WO9739357-A1.
PD 23-OCT-1997.
PF 11-APR-1997; U06049.
PR 12-APR-1996; US-015307.
PA (UJO) UNIV JOHNS HOPKINS.
PA (STRD) UNIV LELAND STANFORD JUNIOR.
PI Andrew D, Bhanot P, Brink M, Hsieh J, Nathans J,
PI Nusse R, Samos CH, Wangy;
DR WPI: 97-526631/48.
DR P-PSDB: W31274.
PT Identification of Wnt receptor binding modulators - useful for
PT treatment of cancer and growth, development or proliferation related
PT disorders
PS Disclosure: Page 46-47; 61pp; English.
CC This sequence comprises novel mouse frizzled gene 8 (Mfz8)
CC that encodes a transmembrane receptor, frizzled-8 (see W31274),
CC a Wnt receptor (WntR). Novel frizzled family members have been
CC identified in Drosophila, mouse, human and Caenorhabditis (see
CC T89885-92) and are considered to be Wnt receptors. Wnt receptors
CC can be used in a novel, claimed method of screening for compounds
CC which modulate the binding of a Wnt polypeptide (secreted proteins
CC involved in cell-to-cell signalling) to a Wnt receptor. Wnt is
CC involved in (mammary) cancer and other processes involving growth,
CC development and proliferation (both normal and abnormal).
CC Modulators identified by the claimed method are useful for

CC treatment of diseases related to these conditions.
SQ Sequence 2421 BP; 403 A; 772 C; 779 G; 467 T;

Query Match 12.18; Score 89; DB 37; Length 2421;
Best Local Similarity 41.18; Pred. No. 3,12e-17;
Matches 124; Conservative 47; Mismatches 131; Indels 0; Gaps 0;

Db 306 tgcgtgtgcaagcagcgtgtaaacacacattacattacacgcccacacagcagc 365
QY 140 TNCARYTNTGTCAGGNATGARTATACARAAATGTMGNTNCCNAAATYNTNGGNCAYG 199
Db 366 acacgcaagatgagcggtgaggtgaccagtttggcgcgtggtgagatagcagt 425
QY 200 ARACNATGAARGATGNTNGARCCNGCNGCNGTGGATNCCNYTNGTATGAARCART 259
Db 426 gctcccccgcacctcaagtctttctgtgtagcatgtacacgcccactcgtcgtgagact 485
QY 260 GYCAVCCNGAYACNAARAARTTYTNTGYWSNTTNTTYGCCNGTNGTGYTTNGAYGAY 319
Db 486 acaagaagcctctgcgcctgtgctctgtgtgtaaacgcgcaagcgcgctgcgcgc 545
QY 320 TNGAYGARACNATNCAACCCCTGTGTCAYWSNTTNTGYTNCARGTNAARGAYMTGTCNC 379
Db 546 cgcctatgcgccagtcagcgttctgtgctgacctgacgcgcgtgcgtgctgctgctg 605
QY 380 CNGTATNGWSNGTNTYGGNTTYCCNTGCCNGAYATGYNGARTGYGAYMGNTTYCCNC 439
Db 606 ag 607
QY 440 AR 441

RESULT 2 standard; DNA; 2334 BP.

ID T89889; AC T89889; DT 27-APR-1998 (first entry); DE Human frizzled gene 5 (Hfz5) encoding a Wnt receptor.
KW Wnt receptor; human frizzled gene 5; Hfz5 gene;
KW signal transduction; cancer; cell growth; cell proliferation;
KW mammary tumour; oncogene; therapy; ds.
OS Homo sapiens. Location/Qualifiers
FH Key 321..2078
FT CDS /*tag= a
FT WO9739357-Al.
PN 23-OCT-1997.
PD 11-APR-1997; U06049.
PR 12-APR-1996; US-015307.
PA (UYJO) UNIV JOHNS HOPKINS.
PI Andrew D, Bhanot P, Brink M, Nathans J,
PI Nusse R, Samos CH, Wangy;
DR WPI: 97-526631/48.
DR P-PSDB: W31271.
PT Identification of Wnt receptor binding modulators - useful for
PT treatment of cancer and growth, development or proliferation related
PT disorders
PS Disclosure; Page 36-37; 6lpp; English.
CC This sequence comprises novel human frizzled gene 5 (Hfz5)
CC that encodes a transmembrane receptor, frizzled-5 (see W31771), a
CC Wnt receptor (WntR). Novel frizzled family members have been
CC identified in Drosophila, mouse, human and Caenorhabditis (see
CC T89885-92) and are considered to be Wnt receptors. Wnt receptors
CC can be used in a novel, claimed method of screening for compounds
CC which modulate the binding of a Wnt polypeptide (secreted proteins
CC involved in cell-to-cell signalling) to a Wnt receptor. Wnt is
CC involved in (mammary) cancer and other processes involving growth,
CC development and proliferation (both normal and abnormal).
CC Modulators identified by the claimed method are useful for
CC treatment of diseases related to these conditions.
SQ Sequence 2334 BP; 356 A; 803 C; 736 G; 439 T;

Query Match 10.28; Score 75; DB 37; Length 2334;

Best Local Similarity 39.7%; Pred. No. 1.24e-12;
Matches 116; Conservative 44; Mismatches 132; Indels 0; Gaps 0;

Db 439 tgtgc 498
QY 146 TTGTCAYCGGNATGARTATACARAAATGTMGNTNCCNAAATYNTNGGNCAYGARACNA 205
Db 499 agacgacgagcgcggtgaggtgaccagttctggtgcgcgtggtgagatccatcgcgc 558
QY 206 TGAARGARTNTNGARCCNGCNGCNGTGGATNCCNYTNGTATGAARCARTGYCAYC 265
Db 559 cggacgc 618
QY 266 CNGAYACNAARAARTTYTNTGYWSNTTNTTYGCCNGTNGTGYTNGAYGAYTNGAYG 325
Db 619 agc 678
QY 326 ARACNATNCAACCCCTGTGTCAYWSNTTNTGYTNCARGTNAARGAYMGNTGYGCCNGTNA 385
Db 679 tgc 730
QY 386 TGWSNGCNTTYGGNTTYCCNTGCCNGAYATGYTNGARTGYGAYMGNTTYCC 437

RESULT 3

ID T89887 standard; DNA; 1770 BP.
AC T89887; DT 27-APR-1998 (first entry); DE Caenorhabditis frizzled gene 1 (Cfz1) encoding a Wnt receptor.
KW Wnt receptor; Caenorhabditis frizzled gene 1; Cfz1 gene;
KW signal transduction; cancer; cell growth; cell proliferation;
KW mammary tumour; oncogene; therapy; ds.
OS Caenorhabditis elegans. Location/Qualifiers
FH Key 57..1634
FT CDS /*tag= a
FT WO9739357-Al.
PN 23-OCT-1997.
PD 11-APR-1997; U06049.
PR 12-APR-1996; US-015307.
PA (UYJO) UNIV JOHNS HOPKINS.
PI (STRD) UNIV LELAND STANFORD JUNIOR.
PI Andrew D, Bhanot P, Brink M, Hsieh J, Nathans J,
PI Nusse R, Samos CH, Wangy;
DR WPI: 97-526631/48.
DR P-PSDB: W31269.
PT Identification of Wnt receptor binding modulators - useful for
PT treatment of cancer and growth, development or proliferation related
PT disorders
PS Disclosure; Page 29-30; 6lpp; English.
CC This sequence comprises novel Caenorhabditis frizzled gene 1 (Cfz1)
CC that encodes a putative transmembrane receptor, frizzled-1 (see
CC W31269), a Wnt receptor (WntR). Novel frizzled family members have
CC been identified in Drosophila, mouse, human and Caenorhabditis (see
CC T89885-92) and are considered to be Wnt receptors. Wnt receptors
CC can be used in a novel, claimed method of screening for compounds
CC which modulate the binding of a Wnt polypeptide (secreted proteins
CC involved in cell-to-cell signalling) to a Wnt receptor. Wnt is
CC involved in (mammary) cancer and other processes involving growth,
CC development and proliferation (both normal and abnormal).
CC Modulators identified by the claimed method are useful for
CC treatment of diseases related to these conditions.
SQ Sequence 1770 BP; 441 A; 438 C; 398 G; 493 T;

Query Match 8.7%; Score 64; DB 37; Length 1770;

Best Local Similarity 37.5%; Pred. No. 4.18e-09;
Matches 118; Conservative 47; Mismatches 150; Indels 0; Gaps 0;

Db 145 atcatgagatgtgcaacgatttgcgtatataacttaacgagcttcccaaatctcgtcgacg 204
QY 137 AYTNCARVTNTGYCAYCGGNATGARTATCARAAATGTMGNTNCCNAAATYNTNGGNC 196
Db 205 aggaatcatggaagagcgctccgcgtccatccatccatccatccatccatccatccatcc 264

[illegible]

```

RESULT      8
ID          NS00031 standard; DNA; 501 BP.
AC          NS00031;
DT          04-SEP-1991 (first entry)
DE          Sequence encoding new modified human beta interferon polypeptides
DE          IFNX 448
DE          Antiviral; cell growth regulator; immune system regulator;
KW          antiproliferative; ss.
OS          Homo sapiens.
FH          key
FT          Location/Qualifiers
FT          1..501
FT          /*tag= a
PN          Ep-163993-A.
PD          11-DEC-1985.
PF          17-MAY-1985; 105750.
PR          17-MAY-1984; GB-012564.
PA          (SEAR ) SEARLE G D & CO.
PI          Bell LD, Boseley PG, Porter AG;
PI          WPI: 85-311944/50.
DR          P-PSDB: P50030.
PT          New modified human beta interferon polypeptide(s) - prepd. by
PT          PT plasmaid transformed bacteria, with improved antiviral,
PT          anti-proliferative and immune regulating actions
PS          Claim 28; Chart 2i, page 40; 71pp; English.
CC          Compared with interferon beta prepd. by recombinant methods, the
CC          INFs of the invention are more active and have different affinities
CC          for cell surface receptors (allowing selective targeting); they
CC          have higher therapeutic index; and improved stability against microbial
CC          breakdown during synthesis; and better in vivo solubility and
CC          stability. They are also easier to recover from incubation mixts.
SQ          Sequence 501 BP; 110 A; 30 C; 59 G; 80 T;

```

[illegible]

RESULT	9	
ID	N50033 standard; DNA; 501 BP.	
AC	N50033;	
DT	04-SEP-1991 (first entry)	
DE	Sequence encoding new modified human beta interferon polypeptides	
DE	INFX 456.	
KW	Antiviral; cell growth regulator; immune system regulator;	
KW	antiproliferative; ss.	
OS	Homo sapiens.	
FT	Key	Location/Qualifiers
FT	cds	1..501
FT		/tag- a
PN	EP-163993-A.	
PD	11-DEC-1985.	
PF	17-MAY-1985; 105750.	
PR	17-MAY-1984; GB-012564.	
PA	(SEAR) SEARLE G D & CO.	
PI	Bell LD, Boseley PG, Porter AG;	
DR	WPI: 85-311944/50.	
DR	P-PSDB: P50032	
PT	New modified human beta interferon polypeptide(s) - prepd. by	
PT	plasmid transformed bacteria, with improved antiviral,	
PT	anti-proliferative and immune regulating actions	
PS	Claim 28; Chart 2k, page 42; 71pp; English.	
CC	Compared with interferon beta prepd. by recombinant methods, the	
CC	INfs of the invention are more active and have different affinities	
CC	for cell surface receptors (allowing selective targeting); they	
CC	have higher therapeutic index; improved stability against microbial	
CC	breakdown during synthesis; and better in vivo solubility and	
CC	stability. They are also easier to recover from incubation mixts.	
SC	Sequence	501 BP; 111 A; 31 C; 68 G; 80 T;

Db 310 garaarytboataargargayttvacnmgdnnaarytbatgwnnnwnnytbcaaytbaar 369
 Cp 283 RRAAYTTTNGTTCNGGRTGRCAYTGYTTTCATNACNARNGNATCCANGCNCNGCYT 224
 Db 370 mgdtayaygngmdathytbcaytaytbaargcnaargartaywnncaytgygcntgg 429
 Cp 223 GYTCNARNACYTCYTTCATNGTTCRTGNCNARNARRTTNGGNARCKCATRTTGTGT 164
 Db 430 acnathgntmgdntngarathytbmgaaytatyttaytathaaaymgdytbacngntay 489
 Cp 163 AYTCAATNCCRTGRCANARYTGNARRTTNGCNGNATNGGTYTTCARTNSWCKYTTT 104
 Db 490 ytbmgdaay 498
 Cp 103 ANSWAART 95

RESULT 10
 ID N50034 standard; DNA; 498 BP.
 AC N50034;
 DT 04-SEP-1991 (first entry)
 DE Sequence encoding new modified human beta interferon polypeptides
 DE IFNX 485.
 KW Antiviral; cell growth regulator; immune system regulator;
 KW antiproliferative; ss.
 OS Homo sapiens.
 FH Key Location/Qualifiers
 FT cds 1..498
 FT /*tag= a
 PN EP-163993-A.
 PD 11-DEC-1985.
 PF 17-MAY-1985; 105750.
 PR 17-MAY-1984; GB-012564.
 PA (SEAR) SEARLE G D & CO.
 PI Bell LD, Boseley PG, Porter AG;
 DR WPI: 85-311944/50.
 DR P-PSDB; P50033.
 PT New modified human beta interferon polypeptide(s) - prepd. by
 PT plasmid transformed bacteria, with improved antiviral,
 PT anti-proliferative and immune regulating actions
 PS Claim 28; Chart 21, page 43; 71pp; English.
 CC Compared with interferon beta prep. by recombinant methods, the
 CC INFs of the invention are more active and have different affinities
 CC for cell surface receptors (allowing selective targeting); they
 CC have higher therapeutic index; improved stability against microbial
 CC breakdown during synthesis; and better in vivo solubility and
 CC stability. They are also easier to recover from incubation mixts.
 SQ Sequence 498 BP; 112 A; 30 C; 68 G; 77 T;

Query Match 7.2%; Score 53; DB 3; Length 498;
 Best Local Similarity 7.8%; Pred. No. 1.09e-05;
 Matches 38; Conservative 183; Mismatches 264; Indels 2; Gaps 2;
 Db 10 aaytybtbgnttytbcarmgdnwnnaaytccarwnncaraarytbybtbgcar 69
 Cp 580 ARTCTIYTRCANARNGTYTCCATNATRCRTTTCRTCTCTCTTTTTRTCANG 521
 Db 70 ytbcaaygngdytbgartaytytbaargaymgdngcngaytayaarathccnatgar 129
 Cp 520 CYTCRCANACYTTTNGGNGCYTCYTCNTGNCNGNARNARRTGTCNSWSNGCNARNG 461
 Db 130 atg-acngaraagarttycaraargaygngcnytbacnathaygaratgytba 188
 Cp 460 GNATRCANARTCTTTTCYTGNGGRANCKRTCRCAITCN-ARCATRTNGGCCANGR 402
 Db 189 raayathtygcnathtymgdcaaygwnnnwnnnacngntggaygaracnathgt 248
 Cp 401 AANCCRAANGCNSWCATNACNGGRCANCKRTCYTTNACYTGNACRCANARNSWRTGR 342
 Db 249 ngaraaytytbcnnaaygntntaycaycarathaaaytbaarcngntnytbarga 308
 Cp 341 CANGGYTGNTNGTYTTCNARTCTCNARRCANACNACNGGRCANARNSWRCANARR 282

Db 309 raarytbgaraargayttvacnmgdnnaarytbatgwnnnwnnytbcaaytbaarmg 368
 Cp 281 AAYTTTNGTTCNGGRTGRCAYTGYTTTCATNACNARNGNATCCANGCNCNGCYTGY 222
 Db 369 dtaytaygngmdathytbcaytaytbaargcnaargartaywnncaytgygcntggac 428
 Cp 221 TCNARNACYTCYTTCATNGTTCRTGNCNARNARRTTNGGNARCKCATRTTGTGTAY 162
 Db 429 nathgntmgdntngarathytbmgaaytatyttaytathaaaymgdytbacngntayt 488
 Cp 161 TCNATNCCRTGRCANARYTGNARRTTNGCNGNATNGGTYTTCARTNSWCKYTTTAN 102
 Db 489 bmgdaay 495
 Cp 101 SWAART 95

RESULT 11
 ID N50028 standard; DNA; 501 BP.
 AC N50028;
 DT 04-SEP-1991 (first entry)
 DE Sequence encoding new modified human beta interferon polypeptides
 DE IFNX 445.
 KW Antiviral; cell growth regulator; immune system regulator;
 KW antiproliferative; ss.
 OS Homo sapiens.
 FH Key Location/Qualifiers
 FT cds 1..501
 FT /*tag= a
 PN EP-163993-A.
 PD 11-DEC-1985.
 PF 17-MAY-1985; 105750.
 PR 17-MAY-1984; GB-012564.
 PA (SEAR) SEARLE G D & CO.
 PI Bell LD, Boseley PG, Porter AG;
 DR WPI: 85-311944/50.
 DR P-PSDB; P50027.
 PT New modified human beta interferon polypeptide(s) - prepd. by
 PT plasmid transformed bacteria, with improved antiviral,
 PT anti-proliferative and immune regulating actions
 PS Claim 28; Chart 2f, page 37; 71pp; English.
 CC Compared with interferon beta prep. by recombinant methods, the
 CC INFs of the invention are more active and have different affinities
 CC for cell surface receptors (allowing selective targeting); they
 CC have higher therapeutic index; improved stability against microbial
 CC breakdown during synthesis; and better in vivo solubility and
 CC stability. They are also easier to recover from incubation mixts.
 SQ Sequence 501 BP; 114 A; 31 C; 68 G; 79 T;

Query Match 7.2%; Score 53; DB 3; Length 501;
 Best Local Similarity 7.7%; Pred. No. 1.09e-05;
 Matches 36; Conservative 180; Mismatches 247; Indels 3; Gaps 2;
 Db 33 dwnwnnaaytccartgcaraatytybtbggcarytbaayggngdytbgartayt 92
 Cp 557 ATWATRTCTRTCTCTCTTTTTRTCANGCYTCRCANACYTTTNGGNGCYTGY 498
 Db 93 ytbcaargaymgdatgaaytygayathcngcngargathaaarcarcngcartyca 152
 Cp 497 TCNGTNGCGNARNARRTGTCNSWSNGCNARNGNATRCANARRCTTTR--TCYT 440
 Db 153 raargaygngcnytbacnathaygaratgytbcaraayathtygcnathtymg 212
 Cp 439 GNGGRAANCKRTCRCAITCN-ARCATRTNGGCCANGGRAANCCRAANGCNSWCATNACN 381
 Db 213 dcargaywnnnwnnnacngntggaygaracnathgtngaraaytytbybcnnaaygt 272
 Cp 380 GNGGRCANCKRTCYTTNACYTGNACRCANARNSWRTGRCANGTYTGNATNGTYTTCN 321
 Db 273 ntaycaycarathaaaytbaarcngntnytbggaraaraaytbgaraargaytt 332
 Cp 320 ARTCTRCNARRCANACNGGRCANARNSWRCANARRAAYTYTITNGTTCNGGRTGR 261

RESULT 14

ID N50023 standard; DNA; 501 BP.
 AC N50023;
 DT 04-SEP-1991 (first entry)
 DE Sequence encoding new modified human beta interferon polypeptides
 DE IFNX 416.
 KW Antiviral; cell growth regulator; immune system regulator;
 KW antiproliferative; ss.
 OS Homo sapiens.
 FH Key Location/Qualifiers
 FT cds 1..501 /*tag= a

EP-163993-A.
 PD 11-DEC-1985. 105750.
 PF 17-MAY-1985; GB-012564.
 PR 17-MAY-1984; SEARLE G D & CO.
 PA (SEAR) SEARLE G D & CO.
 PI Bell LD, Boseley PG, Porter AG;
 DI WPI: 85-311944/50.
 DR P-PSDB; P50022.
 PT New modified human beta interferon polypeptide(s) - prepd. by
 PT plasmid transformed bacteria, with improved antiviral,
 PT anti-proliferative and immune regulating actions
 PS Claim 28; Chart 2a, page 32; 71pp; English.
 CC Compared with interferon beta prepd. by recombinant methods, the
 CC INFs of the invention are more active and have different affinities
 CC for cell surface receptors (allowing selective targetting); they
 CC have higher therapeutic index; improved stability against microbial
 CC breakdown during synthesis; and better in vivo solubility and
 CC stability. They are also easier to recover from incubation mixts.
 CC Sequence 501 BP; 107 A; 31 C; 69 G; 80 T;

Query Match 6.9%; Score 51; DB 3; Length 501;
 Best Local Similarity 16.0%; Pred. No. 4.41e-05;
 Matches 50; Conservative 95; Mismatches 167; Indels 1; Gaps 1;
 Db 166 gcnvthacnathaygaratggtbcaaaathttgcnathttygmddcargaywnwnnn 225
 QY 470 SNWSGAYCAIYNYINCCNGCWACNGARGCCNCCNARGNTGCGARGNTGYAARA 529
 Db 226 wnnacngntggaaygaracnathgtngaraaytytbtgcnaaaygtntaycaycarat 284
 QY 530 AYAAARAAYGAYGAYAAAYGATATNATGARGACNYINTGYAARAAYGAYTYGCTNA 589
 Db 285 haaycsyytbaaracngntytbgarataaarytbgaraargaytytyacnmddgnaa 344
 QY 590 ARATNAARGTNAARGARATNACNTAYATNAAYMGNGAYACNARATNATNTNGARACNA 649
 Db 345 rytbatgwnwnnytbcaaytbaarmgdtaytaygngmdathytbcaytaytbaargc 404
 QY 650 ARWSNAARACNATNAYARIYNAAYGGGTWNSGARGNGAYTYNAARAARWSNGTNY 709
 Db 405 naargattaywnncaytygcntggacnathgtngmdgtngarathytbmddaytyta 464
 QY 710 TNGGVTNAARGAYWSNYINTCARTGYACNTGYGARGARATGAAYGAYATNAAYGCCNT 769
 Db 465 ytyathaaymgd 477
 QY 770 AYTNGTNGATGGG 782

RESULT 15

ID N50032 standard; DNA; 501 BP.
 AC N50032;
 DT 04-SEP-1991 (first entry)
 DE Sequence encoding new modified human beta interferon polypeptides
 DE IFNX 449.
 KW Antiviral; cell growth regulator; immune system regulator;
 KW antiproliferative; ss.
 OS Homo sapiens.
 FH Key Location/Qualifiers

FT cds 1..501 /*tag= a
 FT EP-163993-A.
 PN 11-DEC-1985.
 PD 17-MAY-1985; 105750.
 PF 17-MAY-1984; GB-012564.
 PR (SEAR) SEARLE G D & CO.
 PA Bell LD, Boseley PG, Porter AG;
 PI WPI: 85-311944/50.
 DR P-PSDB; P50031.
 PT New modified human beta interferon polypeptide(s) - prepd. by
 PT plasmid transformed bacteria, with improved antiviral,
 PT anti-proliferative and immune regulating actions
 PS Claim 28; Chart 2j, page 41; 71pp; English.
 CC Compared with interferon beta prepd. by recombinant methods, the
 CC INFs of the invention are more active and have different affinities
 CC for cell surface receptors (allowing selective targetting); they
 CC have higher therapeutic index; improved stability against microbial
 CC breakdown during synthesis; and better in vivo solubility and
 CC stability. They are also easier to recover from incubation mixts.
 CC Sequence 501 BP; 108 A; 30 C; 72 G; 79 T;

Query Match 6.9%; Score 51; DB 3; Length 501;
 Best Local Similarity 7.6%; Pred. No. 4.41e-05;
 Matches 37; Conservative 187; Mismatches 262; Indels 3; Gaps 2;
 Db 10 aaytytbtggnnttytbcarmgdnwnwnnaayttycartgcaraarytbtgtgcar 69
 Cp 580 ARTCTTITTCANARNCTYTCCATNATCTCTTCTCTCTCTCTCTCTCTCTCTCTCT 521
 Db 70 ybaaygngmdybtbartaytytbaargaymgdcaytytgngnttyccngnggar 129
 Cp 520 CYTCRCANACYTTNGGNGCYTCYTCNGTNGCNGGNARNARTRGTCNSWNSWNCNARNG 461
 Db 130 garttygagynaaycarttycaraargaygcngcngnytbacnathaygaratg 189
 Cp 460 GNATRCANARTCTRTT--TCITNGGGRANCKRTCTRCAYTCN-ARCATTCNGGCCANG 404
 Db 130 caraayathtygcgnathtygmddcargaywnwnwnnnacngnttggaaygaracnath 249
 Cp 403 GRAANCCRAANGCNSWCATNACNGGNGCRCANCKRTCTTTCACRCANARNSWRT 344
 Db 250 gtngaraaytytbtgcnaaaygtntaycaycarathaaycaytytbaaracngntytbgar 309
 Cp 343 GRCANGGTGNATNGTYTCRCNARTCTCNARRCANACNGGNGCRAANARNNSWRCANA 284
 Db 310 garaarytbgaraargaytytyacnmddgnaarytbatgwnwnwnnytbcaaytbaar 369
 Cp 283 RKAATTTTNGTFCNGRGTGRCAYTCYTTCATNACNARNGNATCCANGCNCNGCYT 224
 Db 370 mgdtaytaygngmdathytbcaytaytbaargcnaaargartaywnncaytygycntgg 429
 Cp 223 GYTCNARNACYTCYTTCATNGTYTCCTGCCNARNARNTTNGGNARNCCKCATRTTYTGRT 164
 Db 430 acnathgntmgdgtngarathytbmddaytytatythaaymgdytbacngnttay 489
 Cp 163 AYTCNATCCCTGRCANARYTGNARRTTNGCNGGNATNGGTYTTCARTNSWNCYTTT 104
 Db 490 ytbmgdaay 498
 Cp 103 ANSWRAART 95

Search completed: Wed Aug 5 04:19:20 1998
 Job time : 145 secs.

(TM)

Release 3.1A John F. Collins, BioComputing Research Unit.
Copyright (C) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPsrch_nn n.a. - n.a. database search, using Smith-Waterman algorithm

Run on: Wed Aug 5 04:00:48 1998; MasPar time 937.53 Seconds
Tabular output not generated. 1322.069 Million cell updates/sec

Title: >rev-trans-seq-2

Description: (1-885) from translate.seq

Perfect Score: 734

N.A. Sequence: 1 ATGYNCARGGCCGCGGWS.....SNATMGNARYTNCARTGY 885
Comp: TACRANGTYCCGGGCCGCGGWS.....SWANKCNTYRANGTYACR

Scoring table: TABLE default

Gap 6

Nmatch STD : Dbase 0; Query 0

Searched: 1832099 seqs, 700269816 bases x 2

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: emb1-est55

Database: genbank-est107

1: emb1-est10 2: emb1-est11
3: gb_est11 4: gb_est10 5: gb_est11 6: gb_est12 7: gb_est13
8: gb_est14 9: gb_est15 10: gb_est16 11: gb_est17
12: gb_est18 13: gb_est19 14: gb_est20 15: gb_est21
16: gb_est22 17: gb_est23 18: gb_est24 19: gb_est25
20: gb_est26 21: gb_est27 22: gb_est28 23: gb_est29 24: gb_est30
25: gb_est31 26: gb_gss

Statistics: Mean 15.281; Variance 4.528; scale 3.375

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	357	48.6	458	24	AA105749 ml84a03.r1 Stratagene	0.00e+00
2	345	47.0	482	23	W77968 zd70e11.r1 Soares feta	0.00e+00
3	319	43.5	396	25	AA073862 mm97f07.r1 Stratagene	0.00e+00
4	304	41.4	426	20	H87071 ys74d07.r1 Homo sapien	0.00e+00
5	303	41.3	378	10	AA429960 zw58g10.r1 Soares tota	0.00e+00
6	294	40.1	359	4	AA295018 EST100474 Pancreas tum	0.00e+00
7	279	38.0	406	19	AA295018 y128c05.r1 Homo sapien	0.00e+00
8	277	37.7	414	22	AA024771 ze76g10.r1 Soares feta	0.00e+00
9	244	33.2	302	4	AA330758 EST34493 Embryo, 6 wee	8.91e-270
10	242	33.0	344	19	H14917 y126d03.r1 Homo sapien	5.02e-267
11	235	32.0	394	25	AA260087 va52g08.r1 Soares mous	2.08e-257
12	230	28.6	299	22	W08345 mb41f02.r1 Soares mous	3.11e-223
13	193	26.3	248	4	AA297695 EST113252 Fetal heart	3.62e-200

14	176	24.0	219	4	AA297550	EST113078 Fetal kidney	2.85e-177
15	165	22.5	562	7	AA449032	zx07e10.r1 Soares tota	1.45e-162
16	163	22.2	231	19	H29095	ym31e03.r1 Homo sapien	6.71e-160
17	151	20.6	250	23	AA061047	mj78a12.r1 Soares mous	5.48e-144
18	148	20.2	184	24	AA209833	mo90h01.r1 Beddington	4.97e-140
19	145	19.8	523	25	AA122822	mr03c11.r1 Soares mous	4.41e-136
20	144	19.6	188	7	AA457909	vf74b03.r1 Soares mous	9.09e-135
21	143	19.5	477	17	AA934293	ka07g11.s1 Onchocerca	1.87e-133
22	140	19.1	482	18	AA922948	oh10h11.s1 NCI_CGAP-Co	1.60e-129
23	136	18.5	287	24	AA082155	ze88f06.r1 Soares feta	2.70e-124
24	132	18.0	518	17	AA527991	on47d08.s1 NCI_CGAP-Co	4.37e-119
25	128	17.4	334	17	AA834426	of68d09.r1 NCI_CGAP-Co	6.75e-114
26	109	14.9	134	4	AA295743	EST101187 Thymus III H	1.44e-89
27	98	13.4	501	10	AA424593	zv91h11.s1 Soares NHM	8.98e-76
28	92	12.5	520	10	AA424647	zv91h11.r1 Soares NHM	2.32e-68
29	91	12.4	438	10	AA424594	zv91h12.s1 Soares NHM	3.90e-67
30	86	11.7	542	13	AA775798	ad14g08.s1 Soares NBHF	4.83e-61
31	81	11.0	252	13	AA754459	97SN1787 Rice Immature	5.04e-55
32	76	10.4	252	13	AA754459	97SN1787 Rice Immature	4.35e-49
33	72	9.8	337	16	AA907431	o102f01.s1 NCI_CGAP_Lu	2.09e-44
34	72	9.8	369	10	AA423653	ve80h07.r1 Soares mous	2.09e-44
35	66	9.0	424	8	AA481448	zv45b07.s1 Soares ovar	1.63e-37
36	65	8.9	276	5	AA371169	EST83193 Prostate g1an	2.21e-36
37	64	8.7	247	13	AA754458	97SN1784 Rice Immature	2.97e-35
38	63	8.6	531	3	M89402	CEL08B2 Caenorhabditis	3.93e-34
39	61	8.3	247	13	AA754458	97SN1784 Rice Immature	6.63e-32
40	61	8.3	248	5	AA375443	EST87993 HSC172 cells	6.63e-32
41	61	8.3	502	24	AA020088	mh49d08.r1 Soares mous	6.63e-32
42	60	8.2	376	9	C39238	C.elegans cDNA clone y	8.44e-31
43	58	7.9	401	13	AA724275	ah91c05.s1 Soares NFL	1.31e-28
44	57	7.8	360	6	C09584	C.elegans cDNA clone y	1.60e-27
45	56	7.6	360	9	C46856	C.elegans cDNA clone y	1.92e-26

ALIGNMENTS

RESULT	1	AA105749	458 bp	mrna	EST	04-FEB-1997
LOCUS		ml84a03.r1 Stratagene mouse kidney (#937315)				Mus musculus cDNA
DEFINITION		Clone 518668 5' similar to TR.G1151260 G1151260 TRANSMEMBRANE RECEPTOR. ; mRNA sequence.				
ACCESSION		AA105749				
NID		G1654838				
KEYWORDS		EST.				
SOURCE		house mouse.				
ORGANISM		Mus musculus				
REFERENCE		Eukaryotes; mitochondrial eukaryotes; Metazoa; Chordata; Vertebrata; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
AUTHORS		1 (bases 1 to 458) Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T., Geisels,S., Kucaba,I., Lacy,M., Le,M., Martin,J., Morris,M., Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B., Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and Waterston,R.				
TITLE		The WashU-HMI Mouse EST Project				
JOURNAL		Unpublished (1996)				
COMMENT		Contact: Marra M/Mouse EST Project WashU-HMI Mouse EST Project Washington University School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108 Tel: 314 286 1800 Fax: 314 286 1810 Email: mouseest@watson.wustl.edu This clone is available royalty-free through LLNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information. MGI:312516 Seq primer: -28ml3 rev1 ET from Amersham High quality sequence stop: 457. Location/Qualifiers 1. 458 /organism="Mus musculus"				
FEATURES						
source						

```

/strain="C57/BL6"
/notes="vector: pluescript SK-; Site_1: EcoRI; Site_2:
XhoI; Cloned unidirectionally. Primer: Oligo dT. Average
insert size: 1.0 kb; Uni-ZAP XR vector; -5' adaptor
sequence: 5' GAATTCGCGACGAG 3' -3' adaptor sequence: 5'
CTCGAGTTTCTTTTCTTTTCTTTT 3'
/db_xref="taxon:10090"
/clone="518668"
/clone_lib="Stratagene mouse kidney (#937315)"
/sex="females"
/dev_stage="4 weeks"
/lab_host="SOLR (kanamycin resistant)"
<1. >458
mRNA 119 a 140 c 116 g 83 t
ORIGIN
Query Match 48.6%; Score 357; DB 24; Length 458;
Best Local Similarity 60.6%; Pred. No. 0.00e+00;
Matches 280; Conservative 103; Mismatches 75; Indels 4; Gaps 4;

Db 1 AGCAGCGCGGCATGGATTCCGCTGGTCATGAACGACGTCCGCCGACACCAAGAGT 60
Qy 221 ARCARGCNGGNCNTGGATNCNNTYNGTNGTGAACARTGTCATCCNGAYACNARAART 280
Db 61 TCCTGTGCTGCTCTTCGCCCTGTCTCTCGACGACCTAGATGAGACCATCCAGCGGT 120
Qy 281 TTYNTGWSNYNTTYGCNCCNGTCTGYTNGAYGAYTNGAYGARACNATNCARCNT 340
Db 121 GTCACCTGCTCTG-GTGCAGGTGAAGACCGCTGCGGCCGCTCATGTCGCCCTTCG-CT 178
Qy 341 GYCAWWSNYTGTGTCNARGTNAARGAYMGNTGYCCNCCNGTNGATWSNCCNTTYGNT 400
Db 179 TCCCTGCGCCAGACATGCTGAGTGGCGACCTTCCCGCAG-ACAACGACCTCTGCATCC 237
Qy 401 TYCCNTGCCNGAYATGTYNGTGYGAYMGNTTYCCNARGAYAYGAYTNGTYATNC 460
Db 238 CCCTCGCTAGTAG-GACCACCTCTCGCGGCACAGAGAGCTCCCAAGGTGTGTGAAG 296
Qy 461 CNYTNGCWNWSNGAYCAYTYNTNCCNGCNCNARGARGCNCNNAARCTNGYGARG 520
Db 297 CCTGCAAAACCAAGTAATGAGGACGACACGACATCATCGGAACCCCTTGTAAAAATGACT 356
Qy 521 CNTGYAARAAYAAAYAYGAYGAYATNATGABACNNTYNTGYAARAAYGAYT 580
Db 357 TCGCACTGAAATCAAGTGAAGGAGATAACGTATACATCAACAGACACCAAGATCATCC 416
Qy 581 TYGCNTNAARATNARGTNAARGARATNACNTAYATNAAYTNGAYACNARATNATNY 640
Db 417 TGGAGACAAAGACGACCACTTTACAAAGCTGAACGGCGTGT 458
Qy 641 TNGAPACNAARWSNAARACNATNTAYAAARYTNAAYGGNGTWN 682

RESULT 2
LOCUS W77968 462 bp mRNA EST 17-OCT-1996
DEFINITION zd70ell.r1 Soares fetal heart NBHH19W Homo sapiens cDNA clone
sequence 5', mRNA sequence.
ACCESSION W77968
NID 91388502
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryote; mitochondria; eukaryotes; Metazoa; Chordata;
Vertebrata; Euthera; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 462)
AUTHORS Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
Trevisan,E., Waterston,R., Williamson,A., Wohlmann,P. and
Wilson,R.
TITLE The WashU-Merck EST Project
JOURNAL Unpublished (1995)
COMMENT

```

```

Contact: Wilson RK
WashU-Merck EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@wustl.edu
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 1550 Std Error: 0.00
Seq primer: mob.REGA+ET
High quality sequence stop: 424.
FEATURES
Location/Qualifiers
1..462
/organism="Homo sapiens"
/notes="Organ: heart; Vector: pT7T3D (Pharmacia) with a
modified polylinker; Site_1: Not I; Site_2: Eco RI; lstr
strand cDNA was primed with a Not I - oligo(dT) primer [5',
TGTTCACATCTGAAGTGGAGCGCCATCTTTTTTTTTTTT 3'],
double-stranded cDNA was size selected, ligated to Eco RI
adapters (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of a modified pT7T3 vector
(Pharmacia). Library went through one round of
normalization to a Cot = 5. Library constructed by
M.Fatima Bonaldo. This library was constructed from the
same fetus as the fetal lung library, Soares fetal lung
NBHL19W."
/db_xref="taxon:9606"
/clone="346028"
/clone_lib="Soares fetal heart NBHH19W"
/sex="unknown"
/dev_stage="19 weeks"
/lab_host="DH10B (ampicillin resistant)"
<1. >462
mRNA 128 a 133 c 106 g 91 t 4 others
ORIGIN
Query Match 47.0%; Score 345; DB 23; Length 462;
Best Local Similarity 58.0%; Pred. No. 0.00e+00;
Matches 266; Conservative 108; Mismatches 81; Indels 4; Gaps 4;

Db 6 TCGCTTCTCGCCCGCTCTCGCTGATGATGATGACGACGACCATCCAGCCACTCG 65
Qy 289 WSNYNTTYTTCNCCNGTNGTYTNGAYGAYTNGAYGARACNATNCARCCNTGYCAYWSN 348
Db 66 CTCTGN-TGCAGGTGAAGACCGCTGCGCCCGCTCATGTCGGCTTCGN-TTCCCTCTGG 123
Qy 349 YNTGCTNCARGTNAARGAYMGNTGYCCNCCNGTNGTNGNCCNTTYGNTTYCCNTGG 408
Db 124 CCGACATGCTTGAGTCGACCGCTTTCCCGCAGGACACGACCTTTGCATCCCGCTCCG 183
Qy 409 CCNGAYATGTYNGARTGYGAYMGNTTYCCNARGAYAYGAYTNTGYATNCCNTNGCN 468
Db 184 TAGCAGNACCACCTCTCGCCAGCCGACGAGAGCTCCAAAGGTATGTGAAGCTGCAAA 243
Qy 469 WSNWSNGAYCAYTYNTTYCCNCCNARGARGCNCNNAARCTNGYARGCNTGYAAR 528
Db 244 AATAAATAATGATGATGACACGACATATGGAACGCTTTGTAAAAATGATTTTGCACTG 303
Qy 529 AYAARAAYGAYGAYAYATNATGARGACNNTYNTGYARAAYGAYTYGNTYN 588
Db 304 AAAATAAAGTGAAGGAGATAACCTACATCAACCGAGATACCAAAATCATCTCTGGAAGAC 363
Qy 589 AARATNARGTNAARGARATNACNTAYATNAAYMGNGAYACNARATNATNYNGAR-AC 647
Db 364 CAAGAGCAAGACCATTTACAGCTGAACGGTGTGTCGCAAGGAGGACCTTGAAGAAATCGG 423
Qy 648 NAARWSNAARACNATNTAYAAARYTNAAYGGNTGWSNARGMNGAYTYN-AARAARWSNG 706
Db 424 TGCTGTGGCTCAAGACGAGCTTGCACTGCGACCTGTGAAG 462
Qy 707 TNYTNGYTNAARGAYWSNTNCACTGYACNTGYGARG 745

```

```

3
RESULT 3
LOCUS AA073862 396 bp mRNA EST 15-FEB-1997
DEFINITION mm97f07.r1 Stratagene mouse heart (#937316) Mus musculus cDNA clone
536389 5' similar to FR:G1151260 G1151260 TRANSMEMBRANE RECEPTOR.
; mRNA sequence.
ACCESSION AA073862
NID g1595602
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryotes; mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Mus.
REFERENCE 1 (bases 1 to 396)
AUTHORS Marra M., Hillier L., Allen M., Bowles M., Dietrich N., Dubuque T.,
Geisel S., Kucaba T., Lacy M., Le M., Martin J., Morris M.,
Schellenberg K., Streptoe M., Tan F., Underwood K., Moore B.,
Theising B., Wylie T., Lennon G., Soares B., Wilton R. and
Waterston R.
TITLE The WashU-HMI Mouse EST Project
JOURNAL Unpublished (1996)
COMMENT Contact: Marra M/Mouse EST Project
WashU-HMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@wustl.edu
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MG: 323325
Seq primer: -28ml3 rev1 ET from Amersham
High quality sequence stop: 390.
FEATURES
source
1..396
/organism="Mus musculus"
/strain="NIH/Swiss"
/Note="Vector: pBluescript SK; Site_1: EcoRI; Site_2:
XhoI; Cloned unidirectionally. Primer: Oligo dt. 93
pooled NIH/Swiss 13 day embryo hearts. Average insert
size: 1.0 kb; Uni-ZAP XR vector; -5' adaptor sequence: 5'
GAATTCGGCAGCAG 3' -3' adaptor sequence: 5'
CTCGAGTTTCTTTTCTTTT 3'"
/db_xref="taxon:10090"
/clone="536389"
/clone_lib="Stratagene mouse heart (#937316)"
/dev_stage="13 day embryos"
/lab_host="SOLR (kanamycin resistant)"
<1..>396
93 a 126 c 103 g 74 t
mRNA
BASE COUNT 93 a 126 c 103 g 74 t
ORIGIN
Query Match 43.5%; Score 319; DB 25; Length 396;
Best Local Similarity 61.7%; Pred. No. 0.00e+00;
Matches 245; Conservative 88; Mismatches 62; Indels 2; Gaps 2;

Db 1 AGCAGCGGGCGC-TGGATTCCGCTGGTCATGAAGCAGTGCACCGGACACCAAGAAGT 59
1:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 221 ARCARGCNGCNGTGGATNCCNYTNGTNGATGAACARTGYCAYCNGAYACNAARAAT 280
1:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 60 TCCTGTGCTCGCTTCCGCCCTGTCTCTCGAGCAGCTAGATGAGACCATCCAGCGGT 119
1:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 281 TTYTNGYWSNYNTTYGCNCCNGTNGTYTNGAYTNGAYGARACNATNCARCCNT 340
1:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 120 GTCACCTCGCTGTG-CTGAGGTGAGGACCGCTGCAGCCCGCTCATGTCGCTTCGCT 178
1:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 341 GYCAWNSYNTGYTNCARGTNAARGTNGTNGTNGTNGTNGTNGTNGTNGTNGTNGT 400
1:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 179 TCCCTGCGCAGACATGCTGGAGTCGACCGGTTTCCCGCAGGACACGACCTCTGCATCC 238
1:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

```

Db 2  ACGCATNGAATACCAAGATCGGCTGCCCAACCTGCTGGCCAGCAGACCATGAAGG 61
QY 152  AYGCAATNGARTAYCARAAATGNGYTNCCNAAAYNYINGNCAYGARACNATGAAG 211

Db 62  AGTGCTGGAGAGCGCGCGCTGGATCCGCTGGTGCATGAAGCAGTGCACCCGGACA 121
QY 212  ARGNTYNGARCARGCNGGNCNTGGATNCCNTNGTNGATGAARCATGYCAYCCNGAYA 271

Db 122  CCAAGAAGTTCCTGCTGCTCTCCGCCGCCGCTCCGCTCGATGACCTAGACGAGACCA 181
QY 272  CNAARAARTYTYTNGWSNNTTYGNCNCNGTNTGYTNGAYGAYTTNGATGARACNA 331

Db 182  TCCAGCCATGCCACTCGCTGTG-GTGCAGGTGAAAGACGCGCTGCGCCCGGTCTGCC 240
QY 332  TNCARCCTGYCAWSNYTNGYGTNCARGTNAARG-AYWGNITGYGNCNCNGATGWSN 390

Db 241  GCCTTCGGTTCCTCCCTGGCGGACATGCTGTGAGTGCAGCCTGTTCCCCAGGACAGCAG 300
QY 391  GCNTTYGNTTYCCNTGCCNGAYATGYTNGARTGYAYMGNTTYCCNCARGAAYAYGAY 450

Db 301  CTTTGCATCCCCCTNGTAGCAGCAGCACCCTTCTTGTGACGCCAGCAGGAAGTTCAAA 360
QY 451  YTNFYATNCCNTYNGCNSWSNGAYCAIYTTNTNC-CNGCNACNGARGCNCNAA 509

Db 361  GGTATGTAAGCTGNAATAATTAATAATGATGATGAACAACGACATAATGGGAA 414
QY 510  RGNTGYGRCNTGYAARAAYARAAYGAYGAYGA-YAAYGAYATNATGARA 562

RESULT 5
LOCUS AA429960 378 bp mRNA EST 16-OCT-1997
DEFINITION zw58g10.r1 Soares total fetus Nb2HF8 9w Homo sapiens cDNA clone
ACCESSION 774306 5', mRNA sequence.
NID AA429960
KEYWORDS G2113134
SOURCE EST.
ORGANISM human.
Eukaryotae: Metazoa; Chordata; Vertebrata: Mammalia; Eutheria;
Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 378)
AUTHORS Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,
Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin,J.,
Moore,B., Schellenberg,K., Steptoe,M., Tan,F., Theising,B.,
White,Y., Wylie,T., Waterston,R. and Wilson,R.
WashU-Merck EST Project 1997
Unpublished (1997)
Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LNL: contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -28m13 rev2 ET from Amersham
High quality sequence stop: 352.
Location/Qualifiers
1. 378
/organism="Homo sapiens"
/notes="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand
was prepared from mRNA obtained from pooled 8-9 week
(total) fetus material with a Not I - oligo(dT) primer [5'
TGTTACATCTGAATGGAGCGCGCGCTTAAATTTTATTTTATTTT 3'].
Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT73 vector. Library
went through one round of normalization, and was
constructed by Bento Soares and M. Fatima Bonaldo."
/db_xref="taxon:9606"
/clone="774306"

```

```

/clone_lib="Soares total fetus Nb2HF8 9w"
/dev_stage="8-9 weeks"
/lab_host="DH10B"
BASE COUNT 115 a 101 c 88 g 74 t
ORIGIN
Query Match 41.3%; Score 303; DB 10; Length 378;
Best Local Similarity 59.7%; Pred. No. 0.00e+00;
Matches 222; Conservative 89; Mismatches 60; Indels 1; Gaps 1;

Db 8  CCGGTATGTCGCGCTTCG-CTTCCCTCGCCGACATGCTTGAGTGCAGCGTTTCCCC 66
QY 379  CCNGTATGNSGNTTYGNTTYCCNTGCCNGAYATGYTNGARTGYAYMGNTTYCCN 438

Db 67  CAGGACAAGACCTTTGCATCCCTCGCTAGCAGGACCCACTCTCTGCAGCCACCGAG 126
QY 439  CARGAAYAGYATNTGYATNCCNTYNGCNSWSNGAYCAIYTTNTNCNCNACNGAR 498

Db 127  GAAGCTCCCAAGGTATGTGAAGCTGCAAAATAAATAATGATGACACACGACATAATG 186
QY 499  GARGCNCNARGTNTGYGARGCNTGYAARAAYARAAYGAYGAYAYGAYATNATG 558

Db 187  GAAACCTTTGTAATAATGATTTTGCACGTGAATAAATAAGTGAAGGAGATACCTACATC 246
QY 559  GARACNTYTTGYAARAAYGAYTTYGCNTYTAARATNAARGTNAARGARATNACATYATN 618

Db 247  AACCAGATACCAAAATCATCTGGAGACCAAGACCAACCATTTACAGCTGAACGCT 306
QY 619  AAYMGNGATACNARATNATNTNGARACNAAARACNATNTYAAARTTNAAYGN 678

Db 307  GTGTCCGATAGGACCTGAAGAAATCGGTGCTGCTGCTCAAAAGACAGCTTGCAGTCACC 366
QY 679  GTNWSNGMGNGAYTYTAARAARWSNGTNTGTGTTNAARGAYWSNTNCARTGYACN 738

Db 367  TGTGAGGAGATG 378
QY 739  TGYGARGARATG 750

RESULT 6
LOCUS AA295018 359 bp mRNA EST 18-APR-1997
DEFINITION EST100474 Pancreas tumor I Homo sapiens cDNA 5' end, mRNA sequence.
ACCESSION AA295018
NID g1947649
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryotae: Mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae;
Homo.
REFERENCE 1 (bases 1 to 359)
AUTHORS Adams,M.D., Kerlavage,A.R., Fleischmann,R.D., Fuldner,R.A.,
Bult,C.J., Lee,N.H., Kirkness,E.F., Weinstock,K.G., Gocayne,J.D.,
White,O., Sutton,G., Blake,J.A., Brandon,R.C., Man-Wai,C.,
Clayton,R.A., Cline,T.R., Cotton,M.D., Earle-Hughes,J., Fine,L.D.,
Fitzgerald,L.M., Fitchugh,W.M., Fritchman,J.L., Geoghagen,N.S.,
Glodek,A., Gnehm,C.L., Hanna,M.C., Hedblom,E., Hinkle,P.S.Jr.,
Kelley,J.M., Kelley,J.C., Liu,L.-I., Marmaros,S.M., Merrick,J.M.,
Moreno-Palauques,R.F., McDonald,L.A., Nguyen,D.T., Pelligrino,S.M.,
Phillips,C.A., Ryder,E., Scott,J.L., Saudek,D.M., Shirley,R.,
Small,K.V., Spriggs,T.A., Utterback,T.R., Weidman,J.F., Li,X.,
Bednarek,D.P., Cao,L., Cepeda,M.A., Coleman,T.A., Collins,E.J.,
Dimke,D., Feng,D.-F., Ferrie,A., Fischer,C., Hastings,G.A.,
He,W.W., Hu,J.S., Greene,J.M., Gruber,J., Hudson,P., Kim,A.K.,
Kozak,D.L., Kunsch,C., Hungjun,J., Li,H., Weissner,P.S., Olsen,H.,
Raymond,L., Wei,X.F., Wing,J., Xu,C., Yu,G.L., Ruben,S.M.,
Dillon,P.J., Fannon,M.R., Rosen,C.A., Haseltine,W.A., Fields,C.,
Fraser,C.M. and Venter,J.C.
Initial assessment of human gene diversity and expression patterns
based upon 83 million nucleotides of cDNA sequence
Nature 377 (6547 suppl), 3-174 (1995)
JOURNAL 96026280
MEDLINE
COMMENT Other_ESTs: THC180057

```


AUTHORS

Hillier, L., Clark, M., Dubuque, T., Elliston, K., Hawkins, M., Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M., Parsons, J., Rifkin, L., Rohlfing, T., Soares, M., Tan, F., Trevisakis, E., Waterston, R., Williamson, A., Wohlmann, P. and Wilson, R.

TITLE The WashU-Merck EST Project
JOURNAL Unpublished (1995)
COMMENT

Contact: Wilson RK
WashU-Merck EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu

This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -28M13 rev2 from Amersham
High quality sequence stop: 288.

FEATURES

source
1..414
/organism="Homo sapiens"
/note="Organ: heart; Vector: pT73D (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5', TGTTACCAATCTGAAGTGGAGCGGCATCTTTTTTTTTTTT 3'], double-stranded cDNA was size selected, ligated to Eco RI adapters (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified pT73 vector (Pharmacia). Library went through one round of normalization to a Cot = 5. Library constructed by M.Fatima Bonaldo. This library was constructed from the same fetus as the fetal lung library, Soares fetal lung NBHL19W."
/db_xref="taxon:9606"
/clone="364962"
/sex="unknown"
/dev_stage="19 weeks"
/lab_host="DH10B (ampicillin resistant)"
107 a 123 c 95 g 85 t 4 others
BASE COUNT
ORIGIN

Query Match 37.7%; Score 277; DB 22; Length 414;
Best Local Similarity 60.9%; Pred. No. 0.00e+00;
Matches 224; Conservative 85; Mismatches 54; Indels 5; Gaps 5;
Db 1 TGAAGCAGTCCCGGACACCAAGTCTCTGCTCGCTCTTGGCCCGCTGTC 60
QY 251 TGAACARTGCYCCGAYACNAAARATYTYTNTGWSNYNTTYGCNCGNTGY 310
Db 61 TCGATGACCTAGGAGACCATCGCATCGCTCTCGGTGCGAGTGAAGACC 120
QY 311 TNGAYGAYTNGAYGARACNATNCAACNTGTCAYNSNTYTGTCNCAAGTGAAG 370
Db 121 GCTGCGCCCGGTGATGCTGCGCTCGCTGCTGCGCCGGAATGCTGAGTGCAC 180
QY 371 GNTGCGCCNCGTATGWSNCGNTTYGNTTYCCTGCGCNG-AYATGYTNGARTGY 429
Db 181 CGTTTCCCCAGGACACACCTTTGATCCCTCGCTAGCAGCGACACCTCTCCGCA 240
QY 430 MGNTTYCCNARGAYAYGAYTNGATNCCNTGNCNNSNNGAYCAYTYNTNCCN 489
Db 241 GCACCCAGAGACTCCAAGGTATGTGAGCGCTGCNAAATAAATGATGATGACAC 300
QY 490 GCNACNARGARGCNCNCAAGTNTGARGCNTGYAARAAYAAARAAYGAYGAYAA 549
Db 301 GACATTATGGAACGCTTTGTAAAATGATTTTGCACTTGAATAATNAAAGTGAAG 360
QY 550 GAYATNA-TGAGACNCTNTGYAARAAYGAYTYGCTN-ARATN-AARGTNAARG-A 605
Db 361 GATAACCT 368

QY 606 RATNACNT 613
::: ||| |

RESULT

LOCUS 9
DEFINITION AA330758 302 bp mRNA EST 21-APR-1997
ACCESSION EST34493 Embryo, 6 week I Homo sapiens cDNA 5' end, mRNA sequence.
NID AA330758
KEYWORDS G1983000
SOURCE EST.
ORGANISM Homo sapiens
Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homnidae;
Homo.
REFERENCE 1 (bases 1 to 302)
AUTHORS Adams, M.D., Kerlavage, A.R., Fleischmann, R.D., Fuldner, R.A., Bult, C.J., Lee, N.H., Kirkness, E.F., Weinstock, K.G., Gocayne, J.D., White, O., Sutton, G., Blake, J.A., Brandon, R.C., Man, Wai, C., Clayton, R.A., Cline, T.R., Cotton, M.D., Earle-Hughes, J., Fine, L.D., Fitzgerald, L.M., Fitzhugh, W.M., Fritchman, J.L., Geoghagen, N.S., Glodek, A., Gnehm, C.L., Hanna, M.C., Hedblom, E., Hinkle, P.S., Jr., Kelley, J.M., Kelley, J.C., Liu, L.-I., Marmaros, S.M., Merrick, J.M., Moreno-Palauques, R.F., McDonald, L.A., Nguyen, D.T., Pelligrino, S.M., Phillips, C.A., Ryder, S.E., Scott, J.L., Sauder, D.M., Shirley, R., Small, K.V., Spriggs, T.A., Utterback, T.R., Weidman, J.F., Li, Y., Bednarek, D.P., Cao, L., Cepeda, M.A., Coleman, T.A., Collins, E.J., Dimke, D., Feng, D.-F., Ferrie, A., Fischer, C., Hastings, G.A., He, W.W., Hu, J.S., Greene, J.M., Gruber, J., Hudson, P., Kim, A.K., Kozak, D.L., Kunsch, C., Hungjun, J., Li, H., Meissner, P.S., Olsen, H., Raymond, L., Wei, Y.F., Wing, J., Xu, C., Yu, G.L., Ruben, S.M., Dillion, P.J., Fannon, M.R., Rosen, C.A., Haseltine, W.A., Fields, C., Fraser, C.M. and Venter, J.C.
TITLE Initial assessment of human gene diversity and expression patterns based upon 83 million nucleotides of cDNA sequence
JOURNAL Nature 377 (6547 Suppl), 3-174 (1995)
MEDLINE 96026280
COMMENT Other ESTs: THCL80057
Contact: Kerlavage, AR
Bioinformatics
The Institute for Genomic Research
9712 Medical Center Drive, Rockville, MD 20850 USA
Tel: 3018699056
Fax: 3018699423
Email: arkerlavet@igr.org
For clone availability, additional sequence and expression information related to this EST, please check the TIGR Human Gene Index (<http://www.tigr.org/tldb/hgi/hgi.html>)
Seq primer: M13 Reverse.
Location/Qualifiers
1..302
/organism="Homo sapiens"
/note="Organ: embryo; Vector: pBluescript SK-; Site_1: EcoRI; Site_2: XhoI"
/db_xref="ATCC (inhost):132379"
/db_xref="taxon:9606"
/clone.lib="Embryo, 6 week I"
/dev_stage="embryo, 6 wks"
99 a 57 c 82 g 56 t 8 others
BASE COUNT
ORIGIN

Query Match 33.2%; Score 244; DB 4; Length 302;
Best Local Similarity 58.6%; Pred. No. 8.91e-270;
Matches 177; Conservative 69; Mismatches 56; Indels 0; Gaps 0;
Db 1 AACGACATATGAACGCGTTGTAAATAATTTTNCATGAAATAAAGTGAGGAG 60
QY 547 AAYGAYATNATGARACNTNTGYAARAAYGAYTYGCTNNAARATNAARGTNAARG 606
Db 61 ATAACCTACATGACCGGATACCAAAATCATCTTGGAGCAAGCAAGCAACNNTTAC 120
QY 607 ATNACNTATNAAYNGNGAYACNAAATNATNTYNGARACNAAARWNAARACNATNTAY 666


```
/sex="unknown"
/dev_stage="12.5dpc total fetus"
/lab_host="DH10B"
<1..>394
mrna 104 a 96 c 118 g 76 t
BASE COUNT
ORIGIN

Query Match 32.0%; Score 235; DB 25; Length 394;
Best Local Similarity 57.5%; Pred. No. 2,08e-257;
Matches 164; Conservative 71; Mismatches 50; Indels 0; Gaps 0;

Db 1 AAGGAGATACGTACATCAACAGACAGACACCAAGATCATCTCGGAGACAAAGAGAC 60
QY 601 AARGARATNACNATAYATNAAYMGNGAYACNAARATNATNGARACNAARACN 660

Db 61 ATTACAGCTGAACGGCGTGTCCGAAGGACCTGAGAAATCCGTCTGTGGCTCAA 120
QY 661 ATNTAYAAATYNAAYGGNGTNGSNGARNGAYTTNAARASNGTNTYTTGGYTNAAR 720

Db 121 GACAGCTGCAGTCACCTGTGAGGAGATGAACGACATCAACGCTCCGTATCTGGTCA 180
QY 721 GAYNNTNCARTGTACNTGTGARGATGATGAATNAAYGCMCCNTAYTTNGTATG 780

Db 181 GGACAGAACGAGCGGCGAGTGTGTGATCACCCTCCGTGAACCGTGGCAGAGGCGCAG 240
QY 781 GNCARAAACGAGGNGGAGRYTNGTNTATNACNWSNGTNAARMGNTGGCARAAGGNCAR 840

Db 241 AGAGAGTTCAAGCGCATCTCCCGAGAGATCCCGACGCTGCGATGC 285
QY 841 MNGARTTAAARMGNATNWSNMGNSNATNMGNAARYTNCARTGY 885

RESULT 12
LOCUS W08345 299 bp mrna EST 05-SEP-1996
DEFINITION mb41f02.r1 Soares mouse p3NMF19.5 Mus musculus cdna clone 331995
ACCESSION W08345
NID g1282366
KEYWORDS EST.
SOURCE house.
ORGANISM Mus musculus
Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Mus.
REFERENCE 1 (bases 1 to 299)
AUTHORS Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
Geisel,S., Kucaba,I., Lacy,M., Le,M., Martin,J., Morris,M.,
Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
Theising,B., Wylie,T., Lennon,G., Soares,B., Willson,R. and
Waterston,R.
TITLE The WashU-HMI Mouse EST Project
JOURNAL Unpublished (1996)
COMMENT Contact: Marra M/Mouse EST Project
WashU-HMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:213395
Seq primer: mob.REGA+ET
High quality sequence stop: 280.
Location/Qualifiers
1..299
/organism="Mus musculus"
/note="Vector: pT73D (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cdna
was primed with a Not I oligo(dT) primer [5',
TGTTACCAATCGAAGTGGAGCGCGCCATTTTTTTTTTTT 3'],
double-stranded cdna was size selected, ligated to Eco RI
```


TITLE	JOURNAL	MEDLINE	COMMENT
1. <i>Effect of a 10-day course of intravenous cefazolin sodium on the incidence of postoperative wound infection in patients undergoing elective orthopedic surgery</i>	<i>Journal of Clinical Pharmacy and Therapeutics</i>	1990;15:111-114	
2. <i>Effect of a 10-day course of intravenous cefazolin sodium on the incidence of postoperative wound infection in patients undergoing elective orthopedic surgery</i>	<i>Journal of Clinical Pharmacy and Therapeutics</i>	1990;15:111-114	
3. <i>Effect of a 10-day course of intravenous cefazolin sodium on the incidence of postoperative wound infection in patients undergoing elective orthopedic surgery</i>	<i>Journal of Clinical Pharmacy and Therapeutics</i>	1990;15:111-114	
4. <i>Effect of a 10-day course of intravenous cefazolin sodium on the incidence of postoperative wound infection in patients undergoing elective orthopedic surgery</i>	<i>Journal of Clinical Pharmacy and Therapeutics</i>	1990;15:111-114	
5. <i>Effect of a 10-day course of intravenous cefazolin sodium on the incidence of postoperative wound infection in patients undergoing elective orthopedic surgery</i>	<i>Journal of Clinical Pharmacy and Therapeutics</i>	1990;15:111-114	
6. <i>Effect of a 10-day course of intravenous cefazolin sodium on the incidence of postoperative wound infection in patients undergoing elective orthopedic surgery</i>	<i>Journal of Clinical Pharmacy and Therapeutics</i>	1990;15:111-114	
7. <i>Effect of a 10-day course of intravenous cefazolin sodium on the incidence of postoperative wound infection in patients undergoing elective orthopedic surgery</i>	<i>Journal of Clinical Pharmacy and Therapeutics</i>	1990;15:111-114	
8. <i>Effect of a 10-day course of intravenous cefazolin sodium on the incidence of postoperative wound infection in patients undergoing elective orthopedic surgery</i>	<i>Journal of Clinical Pharmacy and Therapeutics</i>	1990;15:111-114	
9. <i>Effect of a 10-day course of intravenous cefazolin sodium on the incidence of postoperative wound infection in patients undergoing elective orthopedic surgery</i>	<i>Journal of Clinical Pharmacy and Therapeutics</i>	1990;15:111-114	
10. <i>Effect of a 10-day course of intravenous cefazolin sodium on the incidence of postoperative wound infection in patients undergoing elective orthopedic surgery</i>	<i>Journal of Clinical Pharmacy and Therapeutics</i>	1990;15:111-114	

Contact: Kerlavage, AR
Bioinformatics
The Institute for Genomic Research
9712 Medical Center Drive, Rockville, MD 20850 USA
Tel: 3018990956
Fax: 3018999423
Email: arkerlav@tigr.org
For clone availability, additional sequence and expression
information related to this EST, please check the TIGR Human Gene
Index (<http://www.tigr.org/tadb/hgi/hgi.html>)
seq primer: M13-21.

```

Index (http://www.tigr.org/tdb/hgi/hgi.html)
Seq primer: M13-21.
Location/Qualifiers
1..219
/organism="Homo sapiens"
/note="Organ: Kidney; Vector: pBluescript KS-; Site_1:
XhoI; Site_2: EcoRI"

```

```

XhoI; Site_2 : ECORT"
/db_xref="ATCC (inhost):142263"
/db_xref="taxon:9606"
/clone_lib="fetal kidney I"
/dev_stage="fetus"
<l .>219
34 a      77 c      60 g      44 t      4 others
          24.0% ; Score 176; DB 4; Length 219;
l Similarity 63.2%; Pred. No. 2.85e-177;
139; Conservative 46; Mismatches 34; Indels 1; Gaps 1;
TTGGAGCAGCGCGGCTTGGATCCCGCTGGTCATGAAGCAGTGCCACCCGGACACCAAG 60
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
TINGARCAGCNGGNCNTGGATNCCNYTNGTNATGAARCACTGCAYCCNGAYACNAAR 276
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
AGTTTCCTGTGCTCTITNGNCCCGCTTGCTTCATGACTAGCAGGACCATTCCAG 120
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
PARTYTYTNGYSWNTNTTGYGCCNCGNTGTGYTNGAYGAYGARACNATNCAR 336

```

CNTGTCAYWSYNTGTGTCNARGTNAARGAYMGNTGCGNCCNGTATGWSNGCNTTY 396
 TGAATTCCTCCGTGGCCGACATGCTTGAGTCGGACCGTTTCC 219
 ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:|
 GNTTTCCTTGGCCNGAYATGTYTNGARTGYCAYMGNTTYC 436
 562 bp mRNA EST 04-JUN-1997
 AA449032 zx07e10.r1 Soares total fetus Nb2HF8 9w Homo sapiens cDNA clone
 785802 5', similar to TR:G310115 G310115 PRIZLED PROTEIN HOMOLOG 2
 PRECURSOR ;, mRNA sequence.
 AA449032
 92163052
 EST.
 human.
 Homo sapiens
 Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
 Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae;
 Homo.
 1 (bases 1 to 562)
 Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,
 Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin,J.,

RESULT	15			EST	04-JUN-1997
LOCUS	AA449032	562 bp	mrna		
DEFINITION	zx07610.r1 Soares total fetus NR2HF8 9w Homo sapiens cDNA clone 785802 5' similar to TR:G310115 G310115 FRIZZLED PROTEIN HOMOLOG 2 PRECURSOR ;, mRNA sequence.				

ACCESSION	AA449032	
NID	g2163052	
KEYWORDS	EST.	
SOURCE	human.	
ORGANISM	Homo sapiens	
	Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;	
	Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae;	
	Homo.	
REFERENCE	1 (bases 1 to 562)	
AUTHORS	Hillier,L., Allten,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,	
	Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin,J.,	

